THIRD EDITION

PRINCIPLES OF TOXICOLOGY

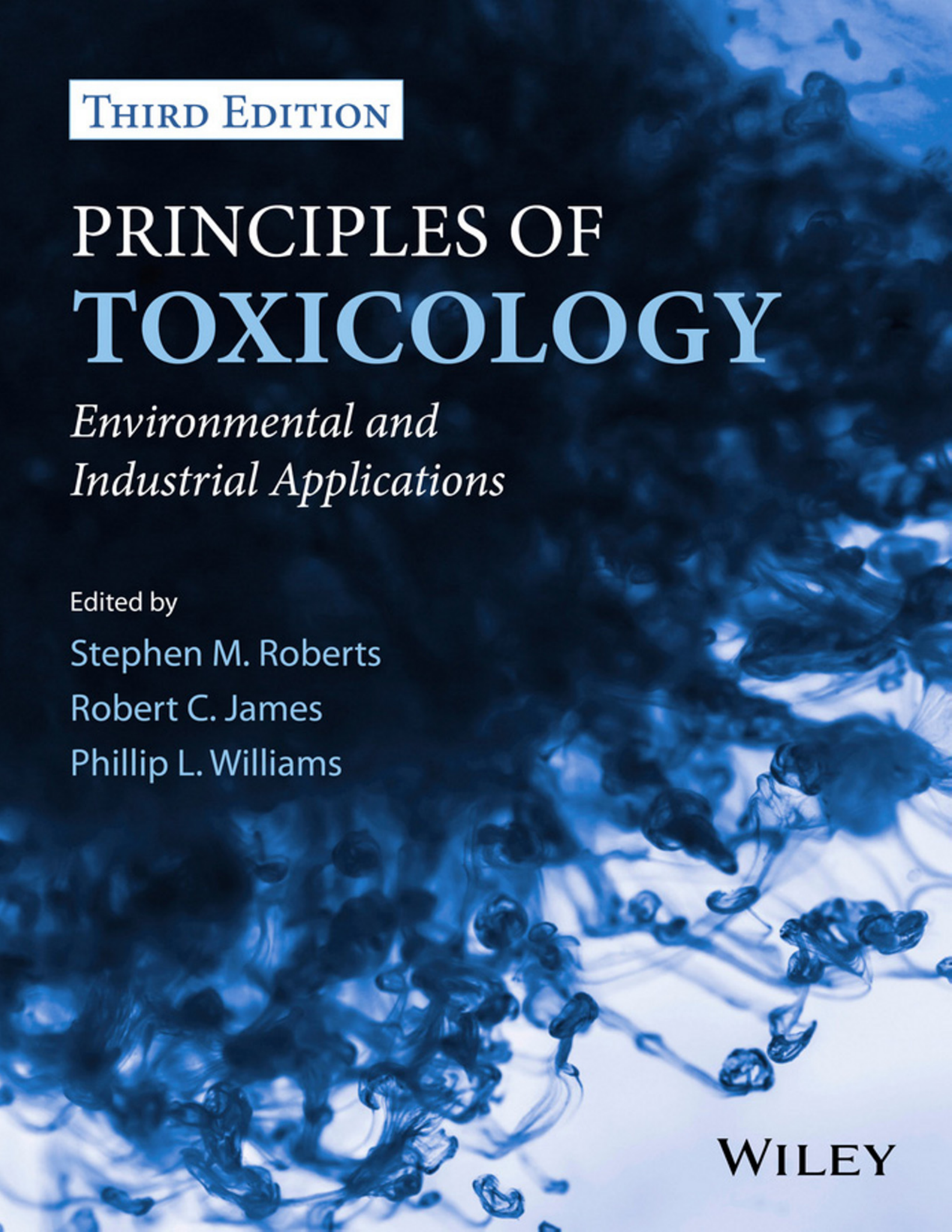
Environmental and Industrial Applications

Edited by

Stephen M. Roberts

Robert C. James

Phillip L. Williams



WILEY

PRINCIPLES OF TOXICOLOGY

PRINCIPLES OF TOXICOLOGY

Environmental and Industrial Applications

Third Edition

Edited By

STEPHEN M. ROBERTS
Professor and Program Director,
Center for Environmental & Human Toxicology,
University of Florida,
Gainesville, FL

ROBERT C. JAMES
Senior Principal Scientist,
ToxStrategies, Inc.,
Boise, ID

PHILLIP L. WILLIAMS
Dean, College of Public Health and Georgia Power Professor of Environmental Health Science,
University of Georgia,
Athens, GA

WILEY

Published by John Wiley & Sons, Inc., Hoboken, New Jersey
Published simultaneously in Canada

Library of Congress Cataloging-in-Publication Data:

Principles of toxicology : environmental and industrial applications / [edited by] Stephen M Roberts, Robert C James, Phillip L Williams. – Third edition.
pages cm
Includes index.
ISBN 978-0-470-90791-7 (cloth)
1. Toxicology. 2. Industrial toxicology. 3. Environmental toxicology. I. Williams, Phillip L., 1952- II. James, Robert C. III. Roberts, Stephen M., 1950-
RA1211.P746 2014
615.9′02–dc23

2014020557

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

CONTENTS

CONTRIBUTORS

T. Renée Anthony, Ph.D., Associate Professor, Department of Occupational and Environmental Health, University of Iowa, University of Iowa, College of Public Health, Iowa City, IA

Lisa A. Bailey, Ph.D., Senior Health Scientist, Gradient, Cambridge, MA

Barbara D. Beck, Ph.D., DABT, FATS, Principal, Gradient, Cambridge, MA

Michael S. Bloom, Ph.D., Associate Professor, Departments of Environmental Health Sciences and Epidemiology and Biostatistics, University at Albany School of Public Health, State University of New York, Rensselaer, NY

Janice Britt, Ph.D., Senior Managing Scientist, ToxStrategies, Inc., Tallahassee, FL

Brenda S. Buikema, M.D., M.P.H, Assistant Professor (Clinical), Department of Occupational and Environmental Health, University of Iowa, College of Public Health; and Division of Pulmonary, Department of Internal Medicine, Critical Care, and Occupational Medicine, University of Iowa, Carver College of Medicine, Iowa City, IA

William M. Caudle, Ph.D., Assistant Professor, Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA

Harvey J. Clewell III, Ph.D., DABT, Director, Center for Human Health Assessment, Hamner Institute for Health Sciences, Research Triangle Park, NC

Rebecca A. Clewell, Ph.D., Hamner Institute for Health Sciences, Research Triangle Park, NC

Mary Jane Cunningham, Ph.D., Founder and President, Nanomics Biosciences, Inc., The Woodlands, TX

Steve Diamond, Ph.D., Vice President, Mid-West Division, Nanosafe, Duluth, MN

Michael R. Franklin, Ph.D., Professor, Department of Pharmacology and Toxicology, University of Utah, Salt Lake City, UT

Michele M. Garber, M.P.H., Senior Scientist, Hazardous Substance & Waste Management Research, Tallahassee, FL

Vincent F. Garry, M.D., MS., DABT, Professor (Emeritus), Department of Laboratory Medicine and Pathology, University of Minnesota, Minneapolis, MN

Fredric Gerr, M.D., Professor of Occupational and Environmental Health and Professor of Epidemiology, University of Iowa, College of Public Health; Iowa City, IA

Sherilyn A. Gross, Ph.D., Senior Health Scientist, Cardno ChemRisk Inc., Boulder, CO

Robert H. Heflich, Ph.D., U.S. Food and Drug Administration, National Center for Toxicological Research, Jefferson, AR

Georgia Hinkley, Ph.D., Research Associate, Center for Environmental & Human Toxicology, University of Florida, Gainesville, FL

Mary L. Hixon, Ph.D., Senior Toxicologist, Gradient Corporation, Cambridge, MA

Keith A. Houck, Ph.D., Principal Investigator, National Center for Computational Toxicology, Office of Research & Development, U.S. Environmental Protection Agency, Research Triangle Park, NC

Robert C. James, Ph.D., Senior Principal Scientist, ToxStrategies, Inc., Boise, ID

Richard S. Judson, Ph.D., Principal Investigator, National Center for Computational Toxicology, Office of Research & Development, U.S. Environmental Protection Agency, Research Triangle Park, NC

James E. Klaunig, Ph.D., Professor, Department of Environmental Health, Indiana University, Ellettsville, IN

Sailesh Konda, M.D., Procedural Dermatology/Mohs Surgery Fellow, Department of Dermatology, Loma Linda University Medical Center, Loma Linda, CA

J. Michael Kuperberg, Ph.D., Center for Biomedical & Toxicological Research, Florida State University, Tallahassee, FL

Lawrence H. Lash, Ph.D., Professor and Associate Chair, Department of Pharmacology, Wayne State University, School of Medicine, Detroit, MI

Ari S. Lewis, M.S., Principal Scientist, Gradient, Cambridge, MA

Cuiqing Liu, Ph.D., Postdoctoral Researcher, Davis Heart and Lung Research Institute, College of Medicine, The Ohio State University, Columbus, OH; and Associate Professor, Department of Physiology, Hangzhou Normal University, Hangzhou, People's Republic of China

Hongbo Ma, Ph.D., Assistant Professor, Zilber School of Public Health, University of Wisconsin – Milwaukee, Milwaukee, WI

Howard I. Maibach, M.D., ATS, Professor, Department of Dermatology, University of California San Francisco, San Francisco, CA

David B. Mayfield, M.S., DABT, Senior Toxicologist, Gradient, Seattle, WA

Gary W. Miller, Ph.D., Professor and Associate Dean for Research, Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA

Martha M. Moore, Ph.D., Senior Manager, ENVIRON International Corp., Little Rock, AR

Meagan B. Myers, Ph.D., Division of Genetic and Molecular Toxicology, National Center for Toxicological Research, Food and Drug Administration, Jefferson, AR

David M. Reif, Ph.D., Associate Professor, Department of Biological Sciences, Bioinformatics Research Center, North Carolina State University, Raleigh, NC

Stephen M. Roberts, Ph.D., Professor and Program Director, Center for Environmental & Human Toxicology, University of Florida, Gainesville, FL

Eric S. Sobel, M.D., Ph.D., Associate Professor and Program Director, Rheumatology and Clinical Immunology, University of Florida College of Medicine, Gainesville, FL

Leah D. Stuchal, Ph.D., Research Assistant Professor, Center for Environmental & Human Toxicology, University of Florida, Gainesville, FL

Qinghua Sun, M.D., Ph.D, Divisions of Environmental Health Sciences & Cardiovascular Medicine, Davis Heart & Lung Research Institute, College of Public Health, The Ohio State University, Columbus, OH

Christopher M. Teaf, Ph.D., Associate Director, Center for Biomedical & Toxicological Research, Florida State University, Tallahassee, FL

Peter L. Truran, Ph.D., Resident Fellow, Center for Philosophy of Science, University of Minnesota, Minneapolis, MN

John E. Vena, Ph.D., Professor and Founding Chair, Department of Public Health Sciences, Medical University of South Carolina, Charleston SC

Phillip L. Williams, Ph.D., Dean, College of Public Health and Georgia Power Professor of Environmental Health Science, University of Georgia, Athens, GA

PREFACE

PURPOSE OF THIS BOOK

Principles of Toxicology: Environmental and Industrial Applications presents compactly and efficiently the scientific basis to toxicology as it applies to the workplace and the environment. The book covers the diverse chemical hazards encountered in the modern work place and natural environment and provides a practical understanding of these hazards for those concerned with protecting the health of humans and ecosystems.

INTENDED AUDIENCE

This book is a third edition and represents an update and expansion on the previous, very successful texts. The first edition of this book was entitled *Industrial Toxicology: Safety and Health Applications in the Workplace.* The current edition retains the emphasis on applied aspects of toxicology, while extending its scope to cover new areas of toxicology such as toxicokinetics, omics technology, nanotoxicology, and computational toxicology. The book was written for those health professionals who need toxicological information and assistance beyond that of an introductory text in general toxicology, yet more practical than that in advanced scientific works on toxicology. In particular, we have in mind industrial hygienists, occupational physicians, safety engineers, environmental health practitioners, occupational health nurses, safety directors, and environmental scientists.

ORGANIZATION OF THE BOOK

This volume consists of 23 chapters. The early chapters establish the scientific basis to toxicology, which is then applied through the rest of the book. It discusses concepts such as absorption, distribution, and elimination of toxic agents from the body. Chapters 5–11 discuss the effects of toxic agents on specific physiological organs or systems, including the blood, liver, kidneys, nerves, skin, lungs, and the immune system.

The next part of the book addresses specific areas of concern in the occupational and environmental settings—both toxic agents and their manifestations. Chapters 12–15 examine the areas of great research interest—reproductive toxicology, developmental toxicology, mutagenesis, and carcinogenesis. Chapters 16–18 examine the toxic effects of metals, pesticides, and organic solvents.

The final part of the book is devoted to specific areas and applications of the toxicological principles from both the environmental and occupational settings. Chapters 19 and 20 cover the emerging areas of nanotoxicology and computational toxicology. Chapters 21 and 22 discuss epidemiologic issues and occupational/environmental health. Chapter 23 covers risk assessment.

FEATURES

The following features from *Principles of Toxicology: Environmental and Industrial Applications* will be especially useful to our readers:

- The book is compact and practical, and the information is structured for easy use by the health professionals in both industry and government.
- The approach is scientific, but applied, rather than theoretical. In this it differs from more general works in toxicology, which fail to emphasize the information pertinent to the industrial environment.
- The book consistently stresses evaluation and control of toxic hazards.

- Numerous illustrations and figures clarify and summarize key points.
- Case histories and examples demonstrate the application of toxicological principles.
- Chapters include suggested reading bibliographies to provide the reader with additional useful information.
- A comprehensive glossary of toxicological terms is included.

STEPHEN M. ROBERTS
ROBERT C. JAMES
PHILLIP L. WILLIAMS

1

GENERAL PRINCIPLES OF TOXICOLOGY

ROBERT C. JAMES, STEPHEN M. ROBERTS, AND PHILLIP L. WILLIAMS

The intent of this chapter is to provide a concise description of the basic principles of toxicology and to illustrate how these principles are used to make reasonable judgments about the potential health hazards and the risks associated with chemical exposures. This chapter explains:

- Some basic definitions and terminology
- What toxicologists study, the scientific disciplines they draw upon, and the specialized areas of interest within toxicology
- Descriptive toxicology and the use of animal studies as the primary basis for hazard identification, the importance of dose, and the generation of dose–response relationships
- How dose–response data might be used to assess safety or risk
- Factors that might alter a chemical's toxicity or the dose–response relationship
- The basic methods for extrapolating dose–response data when developing exposure guidelines of public health interest

1.1 BASIC DEFINITIONS AND TERMINOLOGY

The literal meaning of the term *toxicology* is "the study of poisons." The root word toxic entered the English language around 1655 from the Late Latin word *toxicus* (which meant poisonous), itself derived from *toxikón*, an ancient Greek term for poisons into which arrows were dipped. The early history of toxicology focused on the understanding and uses of different poisons, and perhaps even today most people tend to think of a chemical or products labeled as a "toxic" substance" as that group of chemicals for which minimal exposure inevitably leads to death or some serious long-term adverse effect like cancer. As toxicology has evolved into a modern science it has expanded to encompass all forms of adverse health effects that any substance might produce. The following definitions are provided to help the reader understand several basic terms that may be used in this and other chapters:

Toxic—having the characteristic of being able to produce an undesirable or adverse health effect at some dose.

Toxicity—any toxic (adverse) effect that a chemical or physical agent might produce within a living organism.

Toxicology—the science that deals with the study of the adverse effects (toxicities) that chemicals or physical agents may produce in living organisms under specific conditions of exposure. It is a science that attempts to qualitatively identify all the hazards (i.e., organ toxicities) associated with a substance, as well as to quantitatively determine the exposure conditions under which those hazards/toxicities are induced. Toxicology is the science that experimentally investigates the occurrence, nature, incidence, mechanism, and risk factors for the adverse effects of toxic substances.

As these definitions indicate, the toxic responses that form the study of toxicology span a broad biological and physiological spectrum. Effects of interest may range from something relatively minor such as irritation or tearing to a more serious response like acute and reversible liver or kidney damage, to an even more serious and permanent disability such as cirrhosis of the liver or liver cancer. Given this broad range of potentially adverse effects to consider, it

Principles of Toxicology: Environmental and Industrial Applications, Third Edition. Edited by Stephen M. Roberts, Robert C. James and Phillip L. Williams.

is perhaps useful for those unfamiliar with toxicology to define some additional terms, listed in order of relevance to topics that will be discussed in Chapters 2–24 of this book.

Exposure—a measure of the opportunity for contact with a chemical in one's environment. The presence of a chemical in an environmental media of contact (e.g., in the air we breathe, the water we drink, on surfaces we touch, in foods we might eat). Exposure levels are typically expressed as the concentration of the chemical in the contact medium (e.g., as the ppm concentration in air or water).

Dose—describes the total amount of a toxicant an organism receives as the result of some exposure. The definition of dose typically refers to the *applied dose*, but different definitions and terms arise for the concept of dose as we move from the site of contact on the body to that amount absorbed and then distributed to the various tissues of the body. For example:

Applied dose—this is the total amount of the chemical that is directly applied to or has direct contact with those body surfaces that represent a portal of entry (via absorption) into the body. The applied dose can be higher than the absorbed dose because all of the chemical does not necessarily get across the membranes or surfaces at the site of contact.

Internal/absorbed dose—the actual quantity of a toxicant that is ultimately absorbed into the organism and distributed systemically throughout the body.

Delivered/effective/target organ dose—the amount of toxicant reaching the organ (known as the *target organ*) that is adversely affected by the toxicant.

Acute exposure—exposure that occurs only for a brief period of time (generally <24 h). Often it is considered to be a single exposure (or dose) but may consist of repeated exposures within a short time period.

Subacute exposure—resembles acute exposure except that the exposure duration is greater, for example, from several days to 1 month in animal studies.

Subchronic exposure—exposures repeated or spread over an intermediate time range. For animal testing, this time range is generally considered to be 1–3 months.

Chronic exposure—exposures (either repeated or continuous) over a long period of time. In animal testing this exposure ranges between 90 days to a lifetime. It is generally any exposure that occurs for the majority of that species' lifetime. In occupational settings it is generally considered to be for a number of years or more and may include either a working lifetime or an entire lifetime of an individual.

Acute toxicity—an adverse or undesirable effect that is manifested within a relatively short time interval ranging from almost immediately to within several days following exposure (or dosing). An example would be chemical asphyxiation from exposure to a high concentration of carbon monoxide (CO).

Chronic toxicity—a permanent or lasting adverse effect that is manifested after exposure to a toxicant. An example would be the development of silicosis following a long-term exposure to silica in workplaces such as foundries.

Local toxicity—an adverse or undesirable effect that is manifested at the toxicant's site of contact with the organism. Examples include an acid's ability to cause burning of the eyes, upper respiratory tract irritation, and skin burns.

Systemic toxicity—an adverse or undesirable effect that can be seen anywhere within the organism. It typically involves an organ in the body with selective tissue vulnerability to the toxic effect of the chemical distant from the point of entry of the toxicant (i.e., toxicant requires absorption and distribution within the organism to produce a systemic effect). Examples would include the adverse effects on the kidney or central nervous system (CNS) resulting from the acute or chronic ingestion of mercury.

Reversible toxicity—an adverse or undesirable effect that can be reversed once exposure is stopped. Reversibility of toxicity depends on a number of factors, including the extent of exposure (time and amount of toxicant) and the ability of the affected tissue to repair or regenerate. An example includes hepatic toxicity from acute acetaminophen exposure and liver regeneration.

Delayed or latent toxicity—an adverse or undesirable effect appearing long after the initiation and/or cessation of exposure to the toxicant. An example is cervical cancer during adulthood resulting from in utero exposure to diethylstilbestrol (DES).

Allergic reaction—a reaction to a toxicant caused by an altered state of the normal immune response. The outcome of the exposure can be immediate (anaphylaxis) or delayed (cell-mediated).

Idiosyncratic reaction—a response to a toxicant occurring at exposure levels much lower than those generally required to cause the same effect in most individuals within the population. This response is genetically determined, and a good example would be sensitivity to nitrates due to deficiency in NADH (reduced-form nicotinamide adenine dinucleotide phosphate)–methemoglobin reductase.

Mechanism of toxicity—the necessary biological interactions by which a toxicant exerts its toxic effect on an organism. A simple example is CO asphyxiation due to the binding of CO to hemoglobin, thus preventing the transport of oxygen within the blood.

Toxicant—any substance that causes a harmful (or adverse) effect when in contact with a living organism at a sufficiently high concentration.

Toxin—any toxicant produced by an organism (floral or faunal, including bacteria), that is, naturally produced toxicants. An example would be the pyrethrins, which are natural pesticides produced by pyrethrum flowers (i.e., certain chrysanthemums) that serve as the model for the man-made insecticide class pyrethroids.

Potency—a measure of the ability of a chemical to express its toxicity per unit of dose or dosage. The more potent a chemical, the less dosage needed to induce the toxicity it produces. In general terms, the less potent a chemical is, the safer it is because the probability of achieving a dose sufficient to induce toxicity via a particular route of exposure is lessened. Similarly, more potent chemicals tend to be more dangerous because it takes a smaller dose from an exposure to be able to induced toxicity.

Hazard—the qualitative nature of the adverse or undesirable effect (i.e., the type of adverse effect or toxicity the chemical produces) resulting from exposure to a particular toxicant or physical agent. For example, asphyxiation is the hazard from acute exposures to CO. Cancer, liver toxicity, and immunotoxicity are other hazards (types of toxicities) a chemical exposure might potentially represent. A hazard typically refers to the kind(s) of toxic effect(s) the chemical can produce if the exposure/dose is sufficient.

Safety—the measure or mathematical probability that a specific exposure situation or dose will not produce a toxic effect.

Risk—as generally used in toxicology, the measure or probability that a specific exposure situation or dose will produce a toxic effect.

Risk assessment—the process by which the potential (or probability of) adverse health effects of exposure are characterized. In risk assessment, a safe exposure concentration is extrapolated from the dose–response curve for an adverse effect produced by the chemical that is used to derive a safe exposure concentration. Alternatively, a risk assessment might determine the probability and/or acceptability of a toxicity occurring at a known or measured exposure level.

1.2 TOXICOLOGY: A DIVERSE SCIENCE WITH TWO BASIC GOALS

Toxicology has become a science that builds on and uses knowledge developed in many related medical sciences, such as physiology, biochemistry, pathology, pharmacology, medicine, and epidemiology, to name only a few. Toxicology has evolved from the study of poisons to the study of all adverse effects induced by all chemicals or substances. Although toxicology is a science where a number of areas of specialization have evolved, all toxicologists fall into three principal areas of endeavor: descriptive toxicology, research/mechanistic toxicology, and applied toxicology.

Descriptive toxicologists are scientists whose work focuses on the toxicity testing of chemicals. This work is done primarily at commercial and governmental toxicity testing laboratories, and the studies performed at these facilities are designed to generate basic toxicity information that identifies the various organ toxicities (hazards) the test agent is capable of inducing over those exposure conditions necessary to induce each effect. A thorough description of a chemical's toxicology would identify all possible acute and chronic toxicities, including the genotoxic, reproductive, teratogenic (developmental), and carcinogenic potential of the test agent. It would identify important metabolites of the chemical that are generated as the body attempts to break down and eliminate the chemical, as well as understand how the chemical is absorbed into the body and distributed to tissues throughout the body, identify tissue accumulation or elimination, and ultimately determine how it is excreted from the body. Hopefully, appropriate dose–response test data are generated for those toxicities of greatest concern and that toxicity produced at the lowest dose during the completion of the descriptive studies so that the relative safety of any given exposure or dose level that humans might typically encounter can be predicted.

Basic research or mechanistic toxicologists are scientists who study the chemical or agent in depth for the purpose of gaining an understanding of how the chemical or agent initiates those biochemical or physiological changes within the cell or tissue that result in the toxicity (adverse effect). The goal of mechanistic studies is to understand the specific biological reactions (i.e., the adverse chain of events) within the affected organism that ultimately result in the toxic effect being studied. Mechanistic experiments are performed at the molecular, biochemical, cellular, and tissue level of the affected organism. So, mechanistic assessments may incorporate and apply the knowledge of a number of many other related scientific disciplines within the biological and medical sciences (e.g., physiology, biochemistry, genetics, molecular biology, pathology). Because animal species are generally used to identify chemical-induced hazards, and because there may be significant species-specific responses to a chemical, mechanistic studies help provide the information on those key changes required to induce toxicity, and help reduce the uncertainty of the animal-to-human extrapolation we need to make to develop a safe exposure guideline.

Applied toxicologists are scientists concerned with the use of chemicals in a "real world" or nonlaboratory setting. The primary goal of applied toxicologists is the control of chemical exposures in all work and nonwork environments by setting

safe exposure guidelines for each exposure pathway (e.g., air, skin, ingestion exposure to the chemical) in that environment. Toxicologists who work in this area of toxicology use descriptive and mechanistic toxicity studies to limit the dose received by each or all exposure pathways to a total dose of the chemical that is believed to be safe. The process whereby this safe dose or level of exposure is derived is generally referred to as the area of *risk assessment.* Within applied toxicology a number of subspecialties occur. *Forensic toxicology* is that unique combination of analytical chemistry, pharmacology, and toxicology concerned with the medical and legal aspects of drugs and poisons; it is concerned with the determination of which chemicals are present and responsible in exposure situations of abuse, overdose, poisoning, and death that become of interest to the police, medical examiners, and coroners. *Clinical toxicology* specializes in ways to treat poisoned individuals and focuses on determining and understanding the toxic effects of medicines, simple over-the-counter (nonprescription) drugs, and other household products. *Environmental toxicology* is the subdiscipline concerned with those chemical exposure situations found in our general living environment. These exposures may stem from the agricultural application of chemicals, the release of chemicals during modern-day living (e.g., chemicals released by household products), regulated and unintentional industrial discharges into air or waterways, and various nonpoint emission sources (e.g., the combustion by-products of cars). Within this area there may be even further subspecialization (e.g., ecotoxicology, aquatic toxicology, mammalian toxicology, avian toxicology). *Occupational toxicology* is the subdiscipline concerned with the chemical exposures and diseases found in the workplace, the identification of the hazards or injuries that overexposure to an occupationally used chemical might represent, and the prevention of these exposures or the treatment of the injuries they might produce.

Regardless of the specialization within toxicology, or the types of toxicities of major interest to the toxicologist, essentially every toxicologist performs one or both of the two basic functions of toxicology, which are to (1) examine the nature of the adverse effects produced by a chemical or physical agent (*hazard/toxicity identification* function) and (2) assess the probability of these toxicities occurring under specific conditions of exposure (*dose–response and risk assessment* function). Ultimately, the goal and basic purpose of toxicology is to understand the toxic properties of a chemical so that these adverse effects can be prevented by the development of appropriate handling or exposure guidelines.

1.3 HAZARD IDENTIFICATION FUNCTION

The hazard identification or the discovery of the toxicities a chemical produces requires the testing of chemicals at doses high enough to induce the full spectrum of toxicities a chemical can induce. Typically, the hazard identification process involves traditional animal testing to uncover the spectrum of adverse effects (hazards) the chemical is capable of producing at some dose. One way of characterizing and identifying the hazard is by examining toxicities as a function of exposure duration, as previously described for acute, subacute, subchronic, and chronic exposures.

Because each chemical induces a different spectrum of toxic effects and one does not know beforehand which set of toxicity tests to perform to adequately capture and identify the possible hazards posed by the chemical, the chemical is examined using as wide a range of test systems as possible to ensure that all potential hazards for that chemical have been identified. For a complete toxicological evaluation the typical hazard assessment would follow a scheme similar to that illustrated in Figure 1.1. Typically, one would perform these tests using a tiered approach that starts with short exposure interval testing such as acute and subacute exposure periods (tier 1) and subsequently moves through subchronic tests (tier 2) and then chronic tests (tier 3). At each tier, specialized tests are performed in addition to those assessing target organ toxicities by route of exposure. For example, during the acute testing phase, dermal and reparatory tract irritation may be necessary as well as tests for the development of sensitization by the chemical. During subchronic and chronic testing, target organ testing is augmented by reproductive and developmental studies, testing for immunotoxicity, genotoxicity and mutagenicity, and a chronic bioassay for possible carcinogenic responses.

A tiered approach such as this allows the dose ranges to be set and as the duration of exposure increases, the dose needed to induce the effect is usually lowered (see Table 1.1). The shorter the duration of exposure the lower the cost of the test and the more time-efficient the study. So, trying to identify the end points of interest and toxic dose range is done more time and cost efficiently by seeking the toxicities a chemical induces by testing the chemical short-term tests first. However, both the types of hazards seen and the doses inducing these effects can change with the duration of exposure; and the hazards seen at shorter exposure durations cannot be assumed to be those that will be found after longer durations of exposure. For example, cancer is a latent disease that may require a lifetime of exposure to detect. The route of exposure may also impact the hazard because as the site of absorption is altered it may impact the occurrence of localized effects (like irritation or cellular necrosis at the site of contact) and it can change the tissue distribution as well as the target organ concentration per unit of absorbed dose. Either change may produce a different pattern of target organs affected with different routes of exposure. For example, after testing trichloroethylene (TCE) for carcinogenicity using the mouse as the test organism, it was observed that inhalation exposure induced lung tumors but not liver tumors while

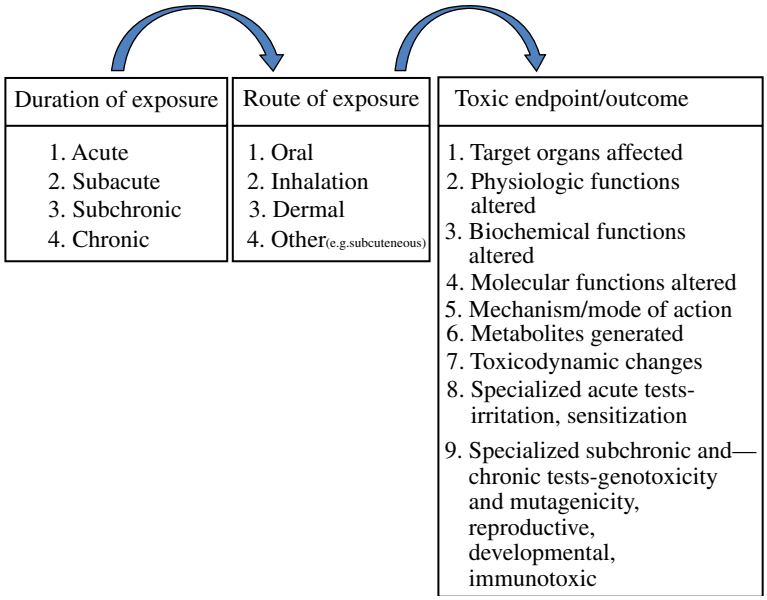


FIGURE 1.1 A generic toxicity testing scheme that shows the ways in which a toxicity test might differ because of the different choices to be made regarding the duration of exposure, the route of exposure, or the endpoint to be measured in the study.

TABLE 1.1 Examples Showing a NOAEL or LOAEL May Change with Exposure Duration

Exposure Duration	Species (Strain)	Organ/End Point	Dose (mg/kg/day)
	a. NOAEL Comparisons		
1,4-Dioxane			
Acute (2 weeks)	Rat (Fischer-344)	Hepatic	1040
Intermediate (13 weeks)			60
Chronic (2 years)			16
Acute (2 weeks)	Rat (Fischer-344)	Renal	1040
Intermediate (13 weeks)			330
Chronic (2 years)			21
Di(2-ethylhexyl)phthalate			
Acute (once)	Rat (Fischer-344)	Renal	5000
Intermediate (90 days)	Rat (Wistar)		1900
Chronic (1 year)	Rat (Sherman)		200
	b. LOAEL Comparisons		
1,4-Dioxane			
Acute (2 weeks)	Rat (Fischer-344)	Hepatic	2750
Intermediate (13 weeks)			150
Chronic (2 years)			81
Acute (2 weeks)	Rat (Fischer-344)	Renal	2750
Intermediate (13 weeks)			760
Chronic (2 years)			103
Di(2-ethylhexyl)phthalate			
Acute (7 days)	Rat (Wistar)	Hepatic	2000
Intermediate (21 days)			1730
Chronic (79 weeks)			1000

oral administration induced liver tumors but not lung tumors. This kind of route-specific toxicity occurs frequently enough that regulatory agencies like the EPA no longer rely upon data gathered by one route of exposure to predict hazards or risk for another route of exposure, that is, there can be considerable uncertainty associated with route-to-route extrapolations without a mechanistic basis for doing so.

Since we are looking for adverse outcomes, the primary source of information for hazard identification comes for toxicity tests using nonhuman species. Over the years, we have developed an extensive array of different toxicity test systems. These test systems are designed to examine end points of interest such as target organs, changes in physiological/biological/molecular function, the different chemical metabolites generated by enzymes whose function is the conversion of both endogenous and exogenous substances into chemical forms more easily eliminated from the body, the mechanism or modes of action, and chemical reactions with key cellular macromolecules (e.g., enzymes, proteins, RNA, DNA).

For example, besides animal or whole organism test results, a toxicologist might use a specialized *in vitro* test system that involves test tube or cell culture methods to examine effects on cellular macromolecules, isolated cell fractions, cellular organelles (e.g., mitochondria), tissue fractions, and isolated perfused whole organs as procedures for examining specific molecular, physiological, or biological functions. A toxicologist might also perform *in vivo* tests in a variety of nonmammalian organisms ranging from simple, single cell organisms (e.g., bacteria, algae) to larger and more complex nonmammalian organisms like nematodes, fruit flies, *Daphnia magna*, or fish, particularly when attempting to identify the ecological hazards or an environmental pollutant.

Some tests are easier and cheaper to perform and can better handle high-volume testing to screen candidate chemicals for further, more detailed toxicity testing or to predict toxicities in chemicals that have not been tested sufficiently via animal tests. One illustration of this approach is where toxicities are receptor-mediated and structure activity relationships may be used as a surrogate measure of subchronic and chronic hazards induced by structurally similar chemicals. The ever-expanding use of *in vitro* test systems may also be desirable in certain situations because they can isolate specific physiological or biochemical pathways in a way that better controls specific test conditions, doses, and outcomes besides being more time- and cost-efficient than whole organism testing. However, *in vitro* tests remove cell or target organism functions from the experimental *in vitro* concentrations (surrogate dose measure) used or the end point being measured may be modified in ways not easily extrapolated to whole organism responses. So, while *in vitro* tests may be undertaken more easily and repeated more consistently, they also have inherently greater uncertainty in comparison to what happens in a whole organism at specific exposure levels or exposure duration. For example, what metabolites are the chemical converted to in whole organisms that are not be seen when using certain *in vitro* test systems? Are toxic or nontoxic metabolites produced by the organism? How does the dose influence the metabolism and distribution throughout the body of the chemical and/or its metabolites? Are the exposure conditions of an *in vitro* system much higher than those that occur in tissues when the chemical is administered in whole animal experiments? In the end, *in vivo* or whole organism testing in a variety of species is generally necessary to identify the range of possible hazards the chemical might pose to humans.

In addition to animal methods, hazard information associated with human exposure to the chemical may also be available. As discussed in more detail elsewhere, there can be significant species differences in the both the beneficial and adverse responses induced by a chemical. So, in the final hazard assessment for a chemical, a toxicologist would like to review as much human data as are available. There are four basic categories of epidemiological information that can assist the hazard evaluation. These categories are occupational epidemiology (mortality and morbidity studies), clinical exposure studies, accidental acute poisonings, and chronic environmental epidemiology studies. The advantages and disadvantages of the hazard information typically provided by these four categories of human toxicological information and that of traditional *in vitro* and animal toxicity tests are summarized and compared in Table 1.2.

1.4 DOSE–RESPONSE/RISK ASSESSMENT FUNCTION

It is probably safe to say that among lay individuals there exists considerable confusion about the term toxic. If asked, most lay individuals would probably define a toxic substance using either a definition that one would apply to highly poisonous or very potently toxic chemicals or something that implies that only some chemicals produce adverse effects in humans and so can be described as toxic chemicals or those substances that we should all avoid. To help illustrate this point, and to begin to emphasize the fact that the toxicity is a function of dose, the reader is invited to take the following pop quiz. First, cross-match the doses shown in column A that produce lethality in 50% of the animals (lethal dose [LD_{50}]) with the chemicals listed in column B. These chemicals are a collection of food additives, medicines, drugs of abuse, poisons, pesticides, and hazardous substances for which the correct LD_{50} is listed somewhere in column A. To perform this cross-matching, first photocopy Table 1.3 and simply mark the ranking of the dose (i.e., the number corresponding next to the dose in column A) you believe correctly corresponds to the chemical it has been measured for in column B. (*Note:* The doses are listed in descending order, and the chemicals have been listed alphabetically. So, the three chemicals you believe to be the safest should have the three largest doses [you should rank them as 1, 2, and 3], and the more unsafe or dangerous you perceive the chemical to be, the higher the numerical ranking you should give it. After testing yourself with the chemicals

TABLE 1.2 Some of the Advantages and Disadvantages of Toxicity Data by Category

Advantages	Disadvantages
a. Occupational Epidemiology (Human) Studies	
May have relevant exposure conditions for the intended use of the chemical.	Exposures (especially past exposures) may have been poorly documented.
As these exposure levels are usually far higher than those found in the general environment, even low or frank effect levels may allow for a realistic extrapolation of a safe level for environmental exposures.	Difficult to properly control; many potential confounding influences (lifestyle, concurrent diseases, genetic, etc.) are inherent to most work populations. These potential confounders are often difficult to identify.
The chance to study the interactive effects of other chemicals that might be present. Again at high doses relative to most environmental situations.	Post facto—not necessarily designed to be protective of health.
	Separating interactive effects resulting from combinations of chemical exposures may be difficult or impossible.
Avoid uncertainties inherent in extrapolating toxicities and dose–response relationships across species.	The increase in disease incidence may have to be large or the measured response severe to be able to demonstrate the existence of the effect being monitored (e.g., cancer). The power to detect risk may be limited.
The full range of human susceptibility (sensitivity) may be measurable if large enough, and diverse enough, populations can be examined.	The full range of human sensitivity for the toxicity of interest may not be measurable because some potentially sensitive populations (young, elderly, infirm) are not represented.
May help identify gender, race, or genetically controlled differences in responses.	Effects must be confirmed by multiple studies as heterogeneous populations are examined and confounders cannot always be excluded.
The potential to study human effects is inherent to almost all industrial uses of chemicals. Thus, a large number of different possible exposure/chemical regimens are available to study.	Often costly and time-consuming. Cost–benefit may be low if confounders or other factors limit the range of exposures, toxicities, confounders, or population variations that might occur with the chemical's toxicity.
b. Clinical (Human) Exposure Studies	
The toxicities identified and the dose–response relationship measured are reported for the most relevant species to study (humans).	The most sensitive group (e.g., young, elderly, infirm) may often be inappropriate for study.
Typically, the components of these studies are better defined and controlled than occupational epidemiology studies. Prospective study design, rather than retrospective design, is used.	Moderately costly to costly to perform.
The chance to study the interactive effects of other chemicals.	Usually limited to shorter exposure intervals than epidemiological studies.
The dose–response relationship is measured in humans. Exposure conditions may be altered during the exposure interval in response to the presence or lack of an effect making NOAELs or LOAELs easier to obtain.	Only NOAELs are targeted for study. These studies are primarily limited to examining safe exposure levels or effects of minimal severity. More serious effects caused by the chemical cannot intentionally be examined by this type of study.
Better than occupational studies for detecting relatively subtle effects. Greater chance to control for the many confounding factors that might be found in occupational studies.	Chronic effects are generally not identifiable by this type of study.
Allows the investigator to test for and identify possible confounders or potential treatments.	Requires study participant compliance.
Allows one to test the specific subpopulations of interest.	May require confirmation by another study.
May help identify gender, race, or genetically controlled differences in responses.	May raise ethical questions about intentionally exposing humans to toxicants.
May be the best method for allowing initial human exposure to the chemical, particularly if medical monitoring is a prominent feature of the study.	Unexpected human toxicities may occur as animal extrapolations are not perfect.
Use of randomization improves the study design and provides best causal inference.	The change being monitored may be statistically significant but still of unknown biological/clinical relevance, leaving the interpretation of results open to question.

(*Continued*)

TABLE 1.2 (Continued)

Advantages	Disadvantages
	c. Environmentally Exposed Epidemiological Studies
The toxicities identified and the dose–response relationship measured are reported for the most relevant species to study (humans).	Exposures to the chemical are typically low relative to other types of human exposures to the chemical in question, or to chemicals causing related toxicities (e.g., exposure to other environmental carcinogens). Thus, attributing the effects observed in a large population may be difficult if many confounding risk factors are present and uncontrolled for in the exposed population.
Exposure conditions are relevant to understanding or preventing significant environmentally caused health effects from occurring.	The exposure of interest may be so low that it is nontoxic and only acting as a surrogate indicator for another risk factor that is present but not identified by the study.
The chance to study the effects of interactive chemicals may be possible.	The number of chemicals with interactive effects may be numerous and their exposures large relative to the chemical of interest. This will confound interpretations of the data.
The full range of human susceptibility may be present.	The full range of human susceptibility may not be present.
May allow one to test specific subpopulations of interest for differences in thresholds, response rates, and other important features of the dose–response relationship.	The full complement of relevant environmental exposure that is associated with the population are not necessarily identified or considered.
May help identify gender, race, or genetically controlled differences in responses.	Large populations may be so heterogeneous in their makeup that when compared to control responses that differences in confounders, gender, age, race, and so on, may weaken the ability to discriminate real disease associations of the chemical exposure from other causes of the disease.
	There may be too many potential confounders to identify and control for and the correlation may be coordinated rather than causal, that is, the problem of the ecological fallacy.
	Exposures are frequently not quantified at the individual level.
	d. Acute Accidental Poisonings
Exposure conditions are realistic for this particular safety extrapolation. In most instances, poisonings are limited to acute exposure situations.	Because the exposure is either accidental or related to a suicide attempt, accurate exposure/dose information is frequently lacking.
These studies often provide a temporal description indicating how the disease will develop in an exposed individual.	This knowledge gained from these studies may be of limited relevance to all other human exposure situations.
Identifies the target organs affected by high, acute exposures. These organs may become candidate targets for chronic toxicity studies.	Confounding factors affecting the magnitude of the response may be difficult to identify as exposure conditions will not be recreated to identify modifying factors.
The clinical response requires no planning as the information gathering typically consists of responding to and treating the organ injuries present as they develop.	Acute toxicities may not mimic those seen with chronic exposure. This may mislead efforts to characterize the effects seen under chronic exposure situations.
	These studies are typically case reports or a small case series and so measures of individual variations in response may be difficult to estimate.
	These chance observations develop without warning, a feature that prevents the development of a systematic study by interested scientists who are knowledgeable about the chemical.
	Because these typically occur as emergency situations, important clinical data may not always be collected.
	e. Animal Toxicity Tests
Easily manipulated and controlled.	Test species response is of uncertain human relevance. Thus, the predictive value is lower than that of human studies.
Best ability to measure subtle responses.	Species/strain/sex/age responses may vary significantly both qualitatively and quantitatively. Thus, a number of different species/strains (both sexes) should ideally be tested.

(Continued)

TABLE 1.2 (Continued)

Advantages	Disadvantages
Widest range of potential toxicities to study.	Exposures levels may not be relevant to (they may far exceed) the human exposure level. The restricted environment of the animal study may not be representative of the complex and variable environment of humans. For example, the practice of allowing animals to eat at will (ad libitum feeding) in bioassays has been shown to increase response rates of certain carcinogens.
Chance to identify and elucidate mechanisms of toxicity that allow for more accurate risk extrapolations to be made using all five categories of toxicity test data.	Selecting the best animal species to study, that is, the species with the most accurate surrogate responses, is always unknown and is difficult to determine a priori (without a certain amount of human test data). Thus, animal data poses somewhat of a Catch-22 situation, that is, you are testing animals to predict human responses to the chemical but must know the human response to that chemical to accurately select the proper animal test species. Mechanisms that are developed may be unique to that species/strain/sex being tested.
Cheaper to perform than full-scale epidemiology studies.	May be a poor measure of the variability inherent to human exposures because animal studies are so well controlled for genetics, doses, observation periods, and so on.
No risk of producing adverse human health effects during the study.	The reproducibility of the animal response may create a false sense of precision when attempting human extrapolations.

Source: Adapted from James et al. (2000).

f. Alternatives to Traditional Animal Testing

Type of Toxicity Test	Advantages	Disadvantages
Structure–activity relationships (SARs)	Does not require the use of any experimental animals. Quick to perform.	Many toxicants with very similar chemical properties have very different toxicities.
In vitro testing	Reduces the number of experimental animals needed. Allows for better control of the toxicant concentration at the target site. Allows for the study of isolated functions such as nerve–muscle interaction and release of neurotransmitter. Easier to control for host factors such as age dependency, nutritional status, and concurrent disease. Possible to use human tissue.	Cannot fully approximate the complexities that take place in whole organisms (i.e., absorption, distribution, biotransformation, and elimination).
Alternative animal testing (nonmammalian and nonavian species)	Less expensive and quicker (due to shorter lifespans) than using higher animals. Since a whole organism is used it allows for absorption, distribution, biotransformation, and elimination of the toxicant.	Since the animal is far removed from humans, the effect of a toxicant can be very different from that found with higher animals.

listed in Tables 1.3, review the correct answers in tables found at the end of this chapter.)

According to the ranking scheme that you selected for these chemicals, were the least potent chemicals common table salt, vitamin K (which is required for normal blood clotting times), the iron supplement dosage added to vitamins for individuals that might be slightly anemic, or a common pain relief medication you can buy at a local drugstore? What were the three most potentially toxic chemicals (most dangerous at the lowest single dose) in your opinion? Were they "natural" or the "synthetic" (human-made) chemicals? How toxic did you rate the nicotine that provides the stimulant properties of tobacco products? How did the potency ranking of prescription medicines like the

TABLE 1.3 Cross-Matching Exercise: Comparative Acutely Lethal Doses

The chemicals listed in this table are *not* correctly matched with their acute median lethal doses (LD_{50}'s). Rearrange the list so that they correctly match. The correct order can be found in the answer table at the end of the chapter.

N	A LD_{50} (mg/kg)	B Toxic Chemical	Correct Order
1	15,000	Alcohol (ethanol)	_____
2	10,000	Arrow poison (curare)	_____
3	4,000	Dioxin or 2,3,7,8-TCDD	_____
4	1,500	(PCBs)—an electrical insulation fluid	_____
5	1,375	Food poison (botulinum toxin)	_____
6	900	Iron supplement (ferrous sulfate)	_____
7	150	Morphine	_____
8	142	Nicotine	_____
9	2	Insecticide (malathion)	_____
10	1	Rat poison (strychnine)	_____
11	0.5	Sedative/sleep aid (phenobarbitol)	_____
12	0.001	Tylenol (acetaminophen)	_____
13	0.00001	Table salt (sodium chloride)	_____

TABLE 1.4 Cross-Matching Exercise: Occupational Exposure Limits—Aspirin and Vegetable Oil Versus Industrial Solvents

The chemicals listed in this table are *not* correctly matched with their allowable workplace exposure levels. Rearrange the list so that they correctly match. The correct order can be found in the answer table at the end of the chapter.

N	Allowable Workplace Exposure Level (mg/m^3)	Chemical (Use)	Correct Order
1	0.05	Aspirin (pain reliever)	_____
2	5	Gasoline (fuel)	_____
3	10	Iodine (antiseptic)	_____
4	54	Perchloroethylene (dry-cleaning fluid)	_____
5	55	Tetrahydrofuran (organic solvent)	_____
6	75	Trichloroethylene (solvent/degreaser)	_____
7	147	1,1,1-Trichloroethane (solvent/degreaser)	_____
8	170	1,1,2-Trichloroethane (solvent/degreaser)	_____
9	890	Toluene (organic solvent)	_____
10	1910	Vegetable oil mists (cooking oil)	_____

sedative phenobarbital or the pain killer morphine compare to the acutely lethal potency of a poison such as strychnine or the pesticide malathion?

Now, take the allowable workplace chronic exposure levels for the following chemicals—aspirin, gasoline, iodine, several different organic solvents, and vegetable oil mists—and again rank these substances going from the highest to lowest allowable workplace air concentration (listed in Table 1.4). Remember that the lower (numerically) the allowable air concentration, the more potently toxic the substance is per unit of exposure. Review the correct answers for tables recreated at the end of this chapter.

Hopefully, the preceding quiz helped illustrate the perceived toxicity or perceived hazard a chemical is thought to pose may mislead one regarding the actual toxic dose or potency of that chemical. As we have defined toxicants (toxic chemicals) as agents capable of producing an adverse effect in a biological system, a reasonable question for one to ask becomes, "Which group of chemicals do we consider to be toxic?" or "Which chemicals do we consider safe?" The short answer to both questions is all chemicals. For even relatively safe chemicals can become toxic if the dose is high enough, and even potent, highly toxic chemicals may be used safely if exposure is kept low enough. As toxicology evolved from the study of substances that were poisonous to a more general study of the adverse effects of all chemicals, the conditions under which chemicals express toxicity became as important as, if not more important than, the kind of adverse effect produced. The importance of understanding the dose at which a chemical becomes toxic (harmful) was recognized centuries ago by Paracelsus (1493–1541), who essentially stated this concept as—"All substances are

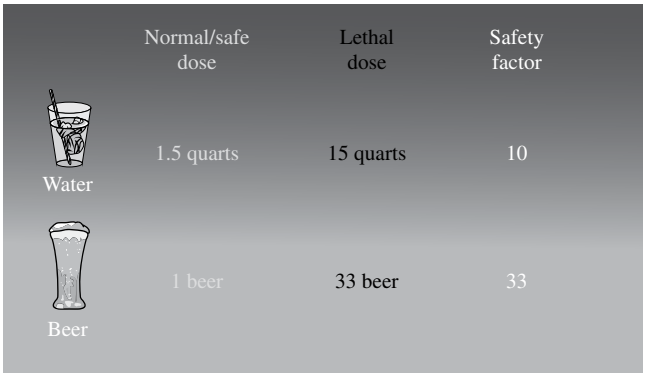


FIGURE 1.2 Acute lethal dose comparisons of two substances commonly used by human populations. *Source:* Adapted from James et al. (2000).

poisons; there is none which is not a poison. The right dose differentiates a poison and a remedy." This statement serves to emphasize the basic functions of toxicology. With the first sentence, Paracelsus tells us that all chemicals express one or more toxicities (hazard identification). However, whether these toxicities are induced or seen is expressed in the second sentence and underscores the second function toxicology—under what dose or exposure conditions is the toxicity expressed. A simple illustration of Paracelsus's admonition and how it applies to all substances is seen Figure 1.2. This figure lists the lethal doses for two substances that most or all adults have been exposed to, water and beer. While some might find it surprising to think that a dose of something as simple and necessary for life as water can be fatal, the ingestion of about 15 quarts of water within a 24-h period is fatal. Normally this toxicity is limited to persons with a serious psychological disorder, but it was also recently illustrated during a radio station–sponsored contest to see who could drink the most water to win a new video game system. One of the contestants vying for the game system unfortunately died the day of the contest from water intoxication. In short, even safe substances are toxic if the dose is high enough. Consequently, another way of viewing the importance of the dose as being key to the toxicity of substances was that provided by Emil Mrak, who sated the concept first attributed to Paracelsus in the following manner—There are no harmless substances, only harmless ways of using substances. An illustration of this principle is exemplified in Figure 1.3 showing that the dose of aspirin increases as one moves through several different desirable target organ effects into those doses that are toxic to other target organs and finally lethality. So, the evaluation of those circumstances under which an adverse effect can be produced is the key to considering whether the exposure is safe or is hazardous. All chemicals are toxic at some dose and may produce harm if the exposure is sufficient (e.g., water or aspirin). Similarly, all chemicals may be used safely under prescribed conditions of dose or usage (e.g., the occupational

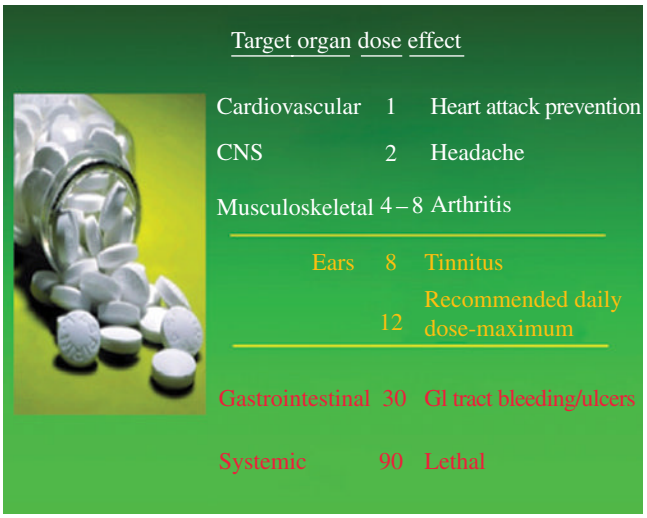


FIGURE 1.3 The dose–response curve for the therapeutic and toxic effects of aspirin.

handling of toxic chemicals during the manufacture of different products). Both quotations serve to remind us that describing a chemical exposure as being either harmless or hazardous is a function of the magnitude of the exposure (dose), and not necessarily the types of toxicities that a chemical might be capable of producing at some dose. Two additional illustrations of this concept are (1) the fact that the vitamins that we consciously take to improve our health and well-being continue to rank as a major cause of accidental poisoning among children; and (2) essentially all the types of toxicities that we associate with the term "hazardous chemicals" are produced by prescription and over-the-counter medication used today. In fact, a number of highly prescribed lipid-lowering drugs produce cancer in certain test animals at high doses but are safely used by many individuals on a daily basis.

Defining Dose and Response

Because all chemicals are toxic at some dose, what judgments determine their use? To answer this, one must first understand the use of the dose–response relationship because this provides the basis for estimating safe and hazardous exposure levels for all chemicals. A dose–response relationship is said to exist when a change in dose produces a consistent, nonrandom change in effect. This effect change can be either in the magnitude of effect or in the percentage of individuals responding at a particular level of effect. For example, the number of animals dying increases as the dose of strychnine is increased, or with therapeutic agents the number of patients recovering from an infection increases as the dosage is increased. In other instances, the severity of the response seen in each animal increases with an increase in dose once the threshold for toxicity has been exceeded.

Dose-Response Graphs

Not only does response to a chemical vary among different species; response also varies within a group of test subjects of the same species. Experience has shown that typically this intraspecies variation follows a normal (Gaussian) distribution when a plot is made relating the frequency of response of the organisms and the magnitude of the response for a given dose. Well-established statistical techniques exist for this distribution and reveal that two-thirds of the test population will exhibit a response within one standard deviation of the mean response, while approximately 95 and 99%, respectively, lie within two and three standard deviations of the mean. Thus, after testing a relatively small number of animals at a specific dose, statistical techniques can be used to define the most probable response (the mean) of that animal species to that dose and the likely range of responses one would see if all animals were tested at that dose (about one or two standard deviations about the mean.) Knowing this for each dose, one can then plot doses, with the standard deviations for each dose, and characterize the dose–response curve and the dose range over which toxicity affects all test organisms (see Figure 1.4).

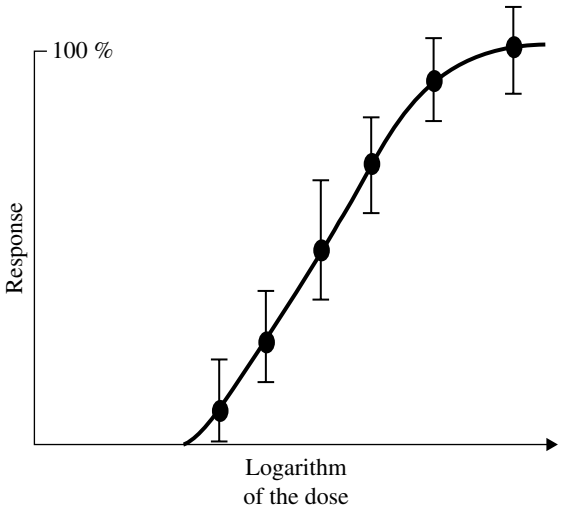


FIGURE 1.4 A simple graphic presentation showing a basic way to portray the dose–response data by plotting the mean responses at each dose and the variation (e.g., standard deviation) about the mean response observed at each dose over the range of doses tested.

In Figure 1.5, a cumulative dose–response curve is featured with a dotted line falling through the highest dose that produces no response in the test animals. Because this dose, and all doses lower than it, fail to produce a toxic response, each of these doses might be referred to as no observable adverse effect levels (NAOELs), which are useful to identify because they represent safe doses of the chemical. The highest of these NAOELs is commonly referred to as the *threshold dose*, which may simply be defined as the dose below which no toxicity is observed (or occurs). For all doses that are larger than the threshold dose, the response increases with an increase in the dose until the dose is high enough to produce a 100% response

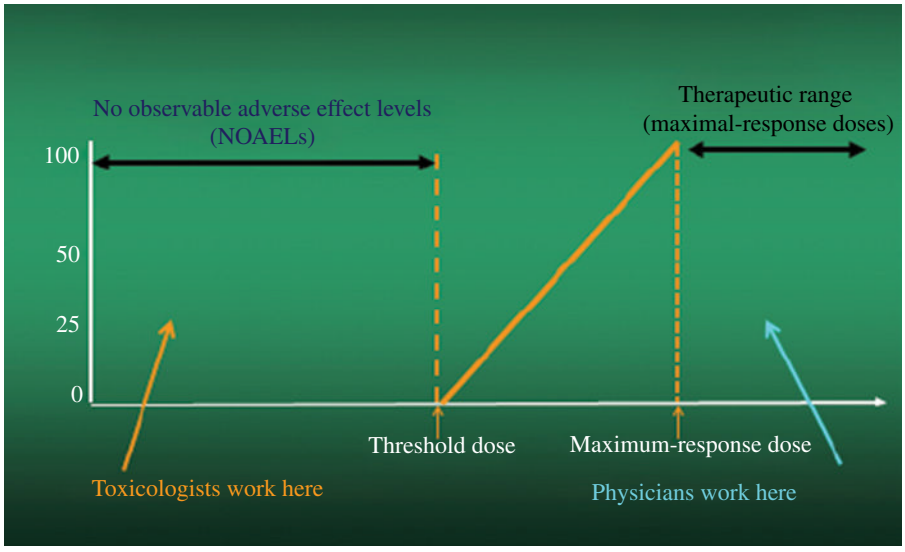


FIGURE 1.5 A schematic representation showing how physicians and toxicologists focus on different responses and areas of the available dose–response curves for a specific chemical. Physicians, because they are interested in producing a beneficial effect from the chemical (drug) in all persons exposed, select those doses in the dose–response range where a maximal response is always achieved. In contrast, toxicologists want to prevent any harmful effects from occurring, and so they select exposures that lead to doses below the threshold of the toxicity so that the harmful response will not occur.

rate (i.e., all subjects respond), and this dose is sometimes referred to as the *maximal-response dose*. All doses larger than the maximal-response dose produce a 100% response, and so the dose–response curve becomes flat again as increasing the dose no longer affects the response rate. For therapeutic effects, this region of the dose–response curve is typically the region physicians seek when they prescribe medicines. Because physicians are seeking a beneficial (therapeutic) effect, typically they would select a dose in this region that is just large enough so that individual variations in response to the dose would still result in a 100% response so as to ensure the efficacy of the drug. In contrast, a toxicologist is generally seeking those doses that produce no response because the effect induced by the chemical is an undesirable one. Thus, toxicologists seek the threshold dose and no-effect region of the dose–response curve.

Before discussing other ways in which dose-response data can be used to assess safety, it will be useful to briefly discuss the various shapes a dose–response curve might take. Although the schematic shape illustrated in Figure 1.6 is the most common shape, the dose–response curve could have either a *supralinear* or *sublinear* shape to it. In Figure 1.6a, the normal linear sigmoid curve is illustrated by line 1; line 2 is an example of a sublinear relationship, and line 3 depicts a supralinear relationship. In addition, some chemicals, while toxic at high doses, produce beneficial effects at low doses. Figure 1.6b–e provides illustrations of the shape of other dose–response relationships. For example, Figure 1.6b depicts the dose–response curve where the doses are not high enough to induce the toxic response being measured. Here no adverse effect is seen regardless of dose. Figure 1.6c depicts a toxicity where the adverse response is a linear function of any dose greater than zero and represents the assumed dose–response relationship that regulatory agencies typically apply to, and model for, carcinogenic substances. Figure 1.6d is a general representation of the most typical dose response curve, the curve for a threshold-dependent toxicity (sometimes referred to as the "hockey stick" dose–response curve), showing that at lower doses the chemical is not capable of inducing an adverse response; then, above a specific dose, toxicity increases as the dose increases.

Figure 1.6e depicts hormesis, which typically has a j-shaped or even a U-shaped curve because at low doses the

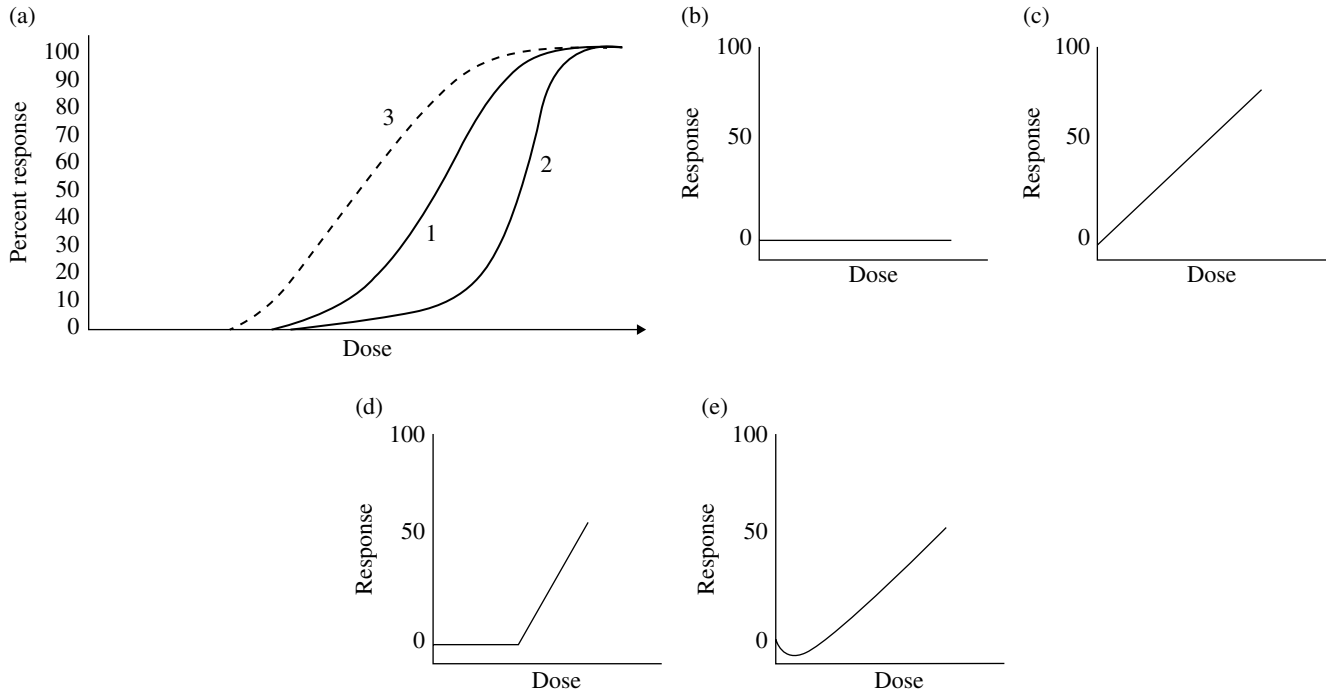


FIGURE 1.6 (a) The dose–response curves with log-linear (1), sublinear (2), and supralinear (3) shapes. (b) The dose–response curve where no effect is seen in the range of doses tested. (c) A graphical depiction of a linear, nonthreshold type of dose–response curve; this shape is typically assumed for carcinogenic substances by regulatory agencies. (d) A graphic representation of a nonlinear, threshold-dependent (toxicity generally seen with noncancer effects; this is commonly referred to as a "hockey stick" shaped dose–response. (e) The "J-shaped" dose–response curve seen with hormesis, a condition where low doses reduce toxicity or represent a beneficial effect that is lost as the dose increases and changes to a toxic responses at even higher doses. Dose–response curves for vitamins, hormones, and medicines frequently express this dose–response curve shape as the desired or beneficial effects are replaced by toxic effects at higher doses.

presence of the chemical benefits the organism and decreases the background response rate of a particular adverse effect. The phenomenon of low-dose stimulation (e.g., growth, reproduction, survival, or longevity) and high-dose inhibition is termed *hormesis*, and the most obvious examples of chemicals that exhibit this phenomenon are vitamins, essential nutrients, and drugs where low doses produce a beneficial effect while higher doses produce toxicity. However, there are other agents that display hormesis for which the benefit of low doses is less intuitive. For example, a number of studies on animals and humans have suggested that low doses of ionizing radiation decrease cancer incidence and mortality, possibly by increasing the presence of DNA repair enzymes, while high doses lead to increased cancer risk. It has been suggested that over time more evidence will show hormesis may be applicable to most, if not all, types of chemical toxicities, but a careful assessment of the extent to which this represents a generalized phenomenon has tended to be hampered by the limited availability of dose–response data below the toxic range for most chemicals. As evidence for hormesis continues to grow, a much clearer understanding of its role will emerge.

1.5 HOW DOSE–RESPONSE DATA CAN BE USED

Dosages are often described as *lethal doses* (LD), where the response being measured is mortality; *toxic doses* (TD), where the response is a serious adverse effect other than lethality; and *sentinel doses* (SD), where the response being measured is a nonadverse or minimally adverse effect. Sentinel effects (e.g., minor irritation, headaches, drowsiness) serve as a warning that greater exposure may result in more serious effects. Construction of the cumulative dose–response curve enables one to identify doses that affect a specific percentage of the exposed population. For example, LD_{50} is the dosage lethal to 50% of the test organisms (see Figure 1.7), or one may choose to identify a less hazardous dose, such as LD_{10} or LD_{01}.

Dose–response data allow the toxicologist to make several useful comparisons or calculations. As Figure 1.7 shows, comparisons of the LD_{50} doses of toxicants A, B, and C indicate the potency (toxicity relative to the dose used) of each chemical. Knowing this difference in potency may allow comparisons among chemicals to determine which is the least toxic per unit of dose (least potent) and therefore the safest of the chemicals for a given dose. This type of comparison may be particularly informative when there is familiarity with at least one of the substances being compared. In this way, the relative human risk or safety of a specific exposure may be approximated by comparing the relative potency of the unknown chemical to the familiar one, and in this manner one may approximate a safe exposure level for humans to the new chemical. For toxic effects, it is typically assumed that humans are as sensitive to the toxicity as the test species. Given this assumption, the test dose producing the response of interest (in units of milligrams per kilogram of body weight (mg/kg)), when multiplied by the average human weight (about 70 kg for a man and 60 kg for a woman), will give an approximation of the toxic human dose.

A relative ranking system developed years ago used this approach to categorize the acute toxicity of a chemical, and is shown in Table 1.5. In this ranking system, the potency of the oral lethal dose of a chemical is used to provide a relative ranking system that characterizes how the toxicity of the chemical is viewed. Again, the least potent category of chemicals (a dose of >15,000 mg/kg for lethality) requires a

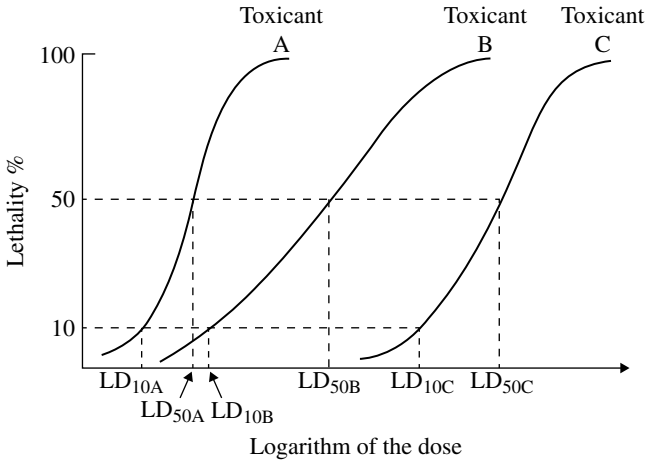


FIGURE 1.7 By plotting the cumulative dose–response curves (log dose), one can identify those doses of a toxicant or toxicants that affect a given percentage of the exposed population. Comparing the values of LD_{50A} to LD_{50B} or LD_{50C} ranks the toxicants according to relative potency for the response monitored.

TABLE 1.5 A Relative Ranking System for Categorization of the Acute Toxicity of a Chemical in Humans

	Probable Oral Lethal Dose	
Toxicity Rating or Class	Dose (mg/kg)	For Average Adult
1. Essentially nontoxic	>15,000	>1 quart
2. Slightly toxic	5,000–15,000	2 cups to 1 quart
3. Moderately toxic	50–5,000	1 ounce to 2 cups
4. Highly toxic	50–500	1 teaspoon to 1 ounce
5. Extremely toxic	1–50	7 drops to 1 teaspoon
6. Supertoxic	<1	<7 drops

Source: Adapted from Canadian Centre for Occupational Health and Safety (CCOHS) (2014).

large oral exposure (e.g., one quart or more) before the substance is lethal. Chemicals like this are considered relatively safe because lethality is unlikely to occur unless a person should ingest a quart or more. As the lethal dose decreases (i.e., becomes more potent at producing lethality), the toxicity rating of the chemical increases because the amount of the dose that may be ingested to incur lethality becomes smaller. Using this ranking system, an industrial hygienist within a work setting might obtain some insight into the acute danger posed by workplace exposure. Similarly, if chronic toxicity is the greatest concern, that is, if the toxicity occurring at the lowest average daily dose is chronic in nature, combining a measure of this toxic dose (e.g., TD_{50}) and appropriate safety factors might generate an acceptable workplace air concentration for the chemical.

Often the dose–response curve for a relatively minor acute toxicity such as odor, tearing, or irritation involves lower doses than more severe toxicities such as coma or liver injury, and much lower doses than fatal exposures. This situation is shown in Figure 1.8, and it can be easily seen that understanding the relationship of the three dose–response curves might allow the use of sentinel effects (represented in Figure 1.8 by the SD curve, the safe dose–response curve) to prevent overexposure and the occurrence of more serious toxicities.

The difference in dose between the toxicity curve and a sentinel effect represents the *margin of safety* (see safety factor expressed in Figure 1.2). In the past, the margin of safety was calculated from data like that shown in Figure 1.8 by dividing TD_{50} by the SD_{50}. This value represents a calculation similar to what physicians refer to as the therapeutic index. The higher the margin of safety, the safer the chemical is to use (i.e., greater room for error). However, if the dose–response curves are complete enough to characterize both the low-response and maximal-response range of doses for both curves then one would generally want to use a more protective definition for the margin of safety (e.g., TD_{01}/SD_{100}). Changing the definition of what the margin of safety represents to include a higher percentile of the sentinel dose–response curve (e.g., the SD_{100}) and correspondingly lower percentile of the toxic dose–response curve (e.g., the TD_{01}) forces the margin of safety to be defined as something protective for the vast majority, if not all, of a population, and so represents an improved method for defining the margin of safety.

Probably the most common use of dose–response data is to use the threshold (or highest NOAEL) dose from an animal toxicity test to extrapolate a corresponding safe dose in humans. Because all exposures producing doses less than the threshold dose (or a NOAEL) should be devoid of toxicity, all exposure below these points will represent safe exposure levels. However, when extrapolating from animal data, as must typically be done in toxicology, there is always some uncertainty as to how closely the animal dose–response data quantitatively and qualitatively mimics the actual human dose–response curve. As a precautionary approach then, safety/uncertainty factors are selected and the NOAEL/threshold dose is divided by a total safety/uncertainty factor from a combination of different uncertainty factors that each reflects the uncertainty of the dose–response data being used in the extrapolation (this is explained in more detail in the risk assessment chapter).

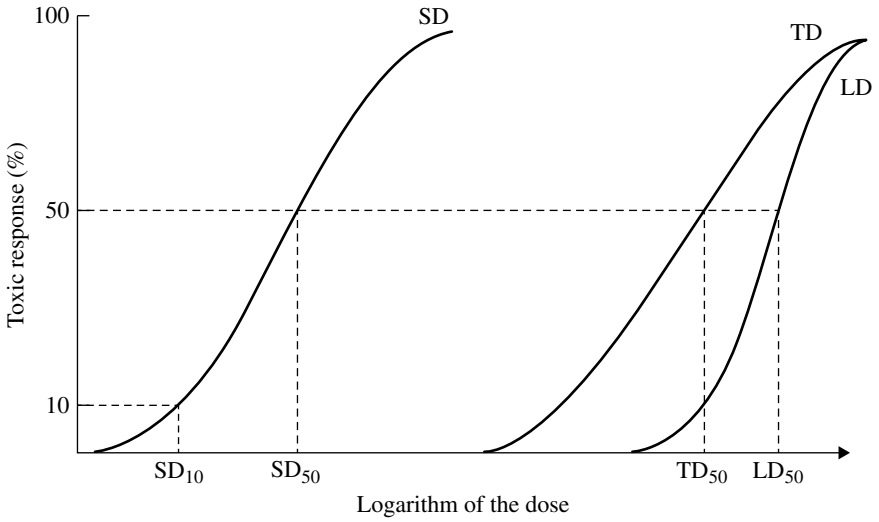


FIGURE 1.8 By plotting or comparing several dose–response curves for a toxicant, one can see the relationship that exists for several responses the chemical might produce. For example, the sentinel response (SD curve) might represent a relatively safe acute toxicity, such as odor or minor irritation to the eyes or nose. The toxic response (TD curve) might represent a serious toxicity, such as organ injury or coma. The lethal response (LD curve), of course, represents the doses producing death. Thus finding symptoms of minor toxicity in a few people at sentinel response (SD_{10}) would be sufficient warning to prevent a serious or hazardous exposure from occurring.

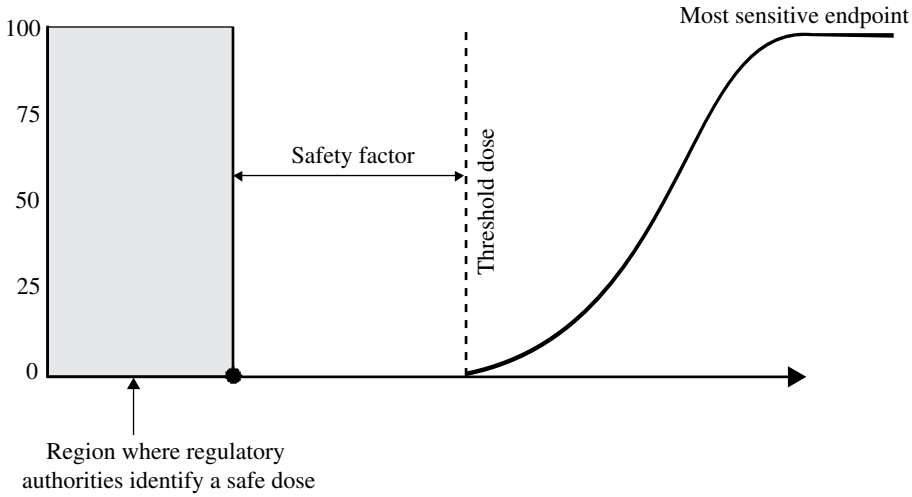


FIGURE 1.9 Graphically demonstrates how the regulatory community develops safe exposure levels for chemical-induced toxicities. Starting at a dose where the adverse effect is not likely by induced (the threshold or NOAEL) dose, the allowable dosage rate is then reduced further by the adoption of safety (uncertainty) factors based on the strength of the available evidence. Reduction by a factor of 10-fold is typically adopted based on the presence of characteristics of the dose–response curve under consideration such as the following: (1) use of animal (nonhuman) data, (2) use of less than chronic exposure duration data, and (3) limited toxicity testing and similar characteristics than contribute to the uncertainty of the extrapolation being made.

After accounting for the potential uncertainty associated with the threshold/NOAEL dose, the final dose selected is considered to be a "safe dosage" that can be used in the development of human exposure guidelines for that chemical (see Figure 1.9). As can be seen in Figure 1.9, the net effect of dividing the threshold or NOAEL dose by some total safety/uncertainty factor is that it is equivalent to selecting a substantially lower dose from the no-effect region of the dose–response curve. This approach essentially adds an additional margin of safety that ensures that the animal data used for the extrapolation has not understated the potency of the chemical in humans.

1.6 AVOIDING INCORRECT CONCLUSIONS FROM DOSE–RESPONSE AND HAZARD IDENTIFICATION DATA

While the dose–response relationship can be determined for each adverse health effect of a toxicant, one must be cognizant of certain limitations when using dose–response data:

1. If only single values from the dose–response curves are available (e.g., the LD_{50}), it must be kept in mind that those values will not provide any information about the shape of the curve. So, while toxicant A in Figure 1.10 would appear to be more toxic than toxicant B chemical at higher doses, this is not true at lower doses. Toxicant B has a lower threshold and actually begins to cause adverse effects at lower doses than toxicant A. Once someone is exposed to a toxicant, the shape of the dose–response curve may be as important as the dose at which toxicity first begins (the threshold dose). Actually, in this regard, toxicant A is a greater concern, not necessarily because of its lower LD_{50} and LD_{100}, but rather because of its steeper dose–response curve. Once individuals become overexposed (exceed the threshold or safe dose), the increase in response occurs with much smaller increases in dose, and more persons are affected with subsequent increases in dose. In other words, once the toxic level is reached, the margin of error for substance A decreases more rapidly than for substance B, because each incremental increase in exposure greatly increases the percentage of individuals affected.
2. Acute toxicity, which is typically generated first because of the savings in time and expense, may not accurately reflect chronic toxicity dose–response relationships. Evidence of this was provided in Table 1.1, which showed the dosage representing the NOAEL for a specific effect decreased as the duration of exposure was increased to subchronic or chronic length of exposure. The type of adverse response generated by a substance may also differ significantly as the exposure duration increases in time and chronic toxicities are sometimes not the same as acute adverse responses. For example, both toluene and benzene cause depression of the CNS, and for this acute effect toluene is the more potently toxic of the two compounds. However, benzene is of greater concern to those with chronic, long-term exposure, because it is carcinogenic while toluene is not. Likewise, the acute hazard of many

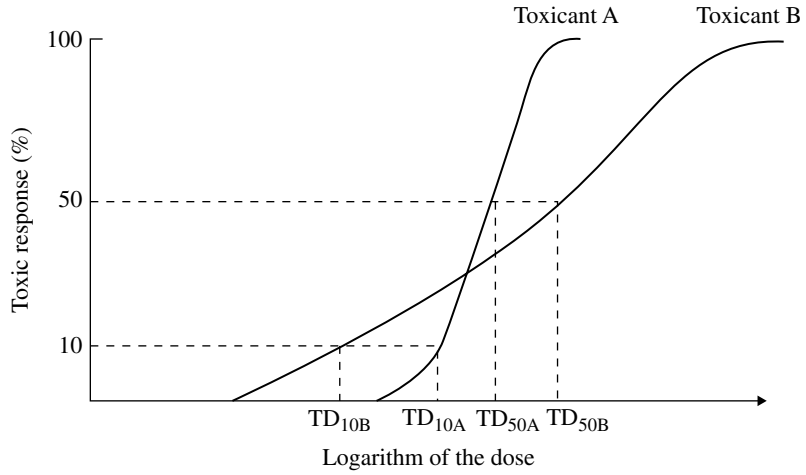


FIGURE 1.10 The shape of the dose–response curve is important. By finding the LD_{50} values for toxicants A and B from a table, one would erroneously assume that A is (always) more toxic than B. The figure demonstrates that this is not true at low doses.

TABLE 1.6 Oral LD_{50} Data for Chloroform

Species	LD_{50} (mg/kg/day)
Rabbit (Dutch Belted)	100[a]
Mouse (CD-1)	250
Human	602
Rat (Sprague-Dawley)	908
Mouse (Swiss)	1100
Mouse (ICR-Swiss)	1400[b]
Rat (Wistar)	2180

Source: Adapted from ATSDR (1996), *Toxicant Profile for Chloroform*.
[a]Based on 13 days of dosing.
[b]Female mice.

TABLE 1.7 Chloroform Toxicity: Inhalation Studies

Mouse as Test Species	Exposure (ppm)	Rat as Test Species
NOAEL—liver	3	NOAEL—respiratory tract
LOAEL—liver	10	LOAEL—respiratory tract (nasal)
	10	NOAEL—kidney
	30	LOAEL—kidney
AEL—liver	100	
NOAEL—kidney	100	NOAEL—liver
LOAEL—kidney	300	LOAEL—liver
NOAEL—respiratory tract	300	

Source: Adapted from ATSDR (1996), *Toxicant Profile for Chloroform.* The duration of exposure for all tests were 6 h/day for 7 days. NOAEL (no observable adverse effect level), LOAEL (low observable effect level—mild organ injury), AEL (adverse effect level—severe/frank organ injury).

chlorinated solvents is generally limited to the CNS depressant properties of the chemical, while for chronic exposure liver and/or kidney effects and possibly cancer may become the primary concern.

3. There is usually little information for guidance in deciding what animal data will best mimic the human response. For example, a question that often arises initially in the study of a chemical is the following: is the test species less sensitive or more sensitive than humans? As shown in Table 1.6, the dose of chloroform that is lethal to 50% of the test animals (i.e., the LD_{50}) varies depending on the species and strain of animal tested. Estimation of the fatal human dose based on the animal results (shown in Table 1.6) would overstate the toxicity of chloroform when using the rabbit or CD-1 mouse data, and underestimate the toxicity of chloroform if projecting lethality using data from the two remaining mouse strains or the two rat strains tested. Another example was illustrated in Table 1.7, where the pattern of toxicity as exposure increased was significantly different when comparing results for the mouse versus those for the rat. In short, dose–response curves for a species less sensitive to the toxicity of interest than humans will understate or underpredict the potential for harm (the risk), while dose–response curves for a species more sensitive or uniquely sensitive (incurs a toxicity not seen in other species) to a specific toxicity will overstate or exaggerate the potential for harm (the risk). The only way to know which species provides the most accurate dose and hazard extrapolations is if you have reliable mechanistic data to be able to pick the correct species to use.
4. In subchronic and chronic testing, generally the highest dose to test is defined as the *maximally tolerated dose* (MTD). This dose is generally defined as the highest dose the test species can be given without generating excessive systemic toxicity (it is usually defined as a dose not inducing a greater than 10%

decrease in body weight or sufficient lethality that adversely impacts completion of the test). Use of the MTD in subchronic and chronic animal tests represents a compromise between two desired goals, sensitivity (avoiding false negatives) and specificity (avoiding false positives). The first goal is to ensure all possible hazards of the chemical have been identified. Here, testing the highest dose possible is desirable because it reduces the chance a false negative will be generated simply because the doses tested were too low to generate an observable response in a study with an adequate number of test animals. In addition, the higher the incidence of the response (usually a function of dose), the smaller the number of animals that have to be used to be able to see a statistically significant change. So, using the MTD increases the ease of observing a positive response as well as reduces cost and space needs, which in turn increases the ability to test more chemicals over a shorter interval of time. For these two reasons, testing the highest dose possible is desirable, especially for regulatory purposes. The second desired goal is that the test paradigm not be one that generates a high percentage of false positives, and thereby potentially eliminate or severely restrict the use of chemicals that may benefit society. However, the use of very high doses of some chemicals can create cellular, bio-chemical, and physiological changes that can produce chronic organ toxicity or carcinogenicity, but do so only under these altered cellular conditions created by the high dose. Where lower doses do not induce these cellular or biochemical changes toxicity is not induced or expected. For example, Gold noted that 44% of the chemical carcinogens she reviewed were not capable of inducing a carcinogenic response at doses as high as ¼–½ the MTD. So, while the MTD might induce all possible toxicities and identify the complete list of potential human hazards, concern for and consideration of these hazards may exaggerate the potential human harm if they are only seen at exposures and doses that substantially exceed the worst-case human exposure scenario.

Species-Related Differences in Hazard and Dose–Response Information: A Frequent Problem that Raises Uncertainty for Animal-to-Human Extrapolations

The basic premise for using animal data to try and predict safe or unsafe human exposure situations is that the effects seen in animal tests are applicable to, and predictive of, the human response. That is, the premise of animal toxicity testing is that the animal response is qualitatively and quantitatively the same, or very similar to, that found in humans. Unfortunately, species-specific differences in the pattern of toxicities observed or in the potency of the chemical are a relatively common phenomenon in toxicology. There are numerous reasons for these species-specific differences, but in the end one can state that there are genetically controlled differences among species that produce anatomical, physiological, biological, and biochemical differences across animal species, or between rodent species (the most frequently used test species in toxicity testing) and man. These differences may confound the animal-to-human extrapolation by increasing the uncertainty and concern we have for the accuracy of both the hazard extrapolation and the dose–response extrapolation being made.

For example, some laboratory animals possess certain anatomical features that humans lack, such as the Zymbal gland and a forestomach. So, when a chemical produces organ toxicity or cancer within these structures, the human relevance for such findings is unknown. Similarly, male rats produce a protein known as α-2microglobulin, which has been shown to interact with the metabolites of certain chemicals in a manner that results in repeated cellular injury within the kidney. This reaction is believed to be responsible for the kidney tumors seen in the male rat after chronic exposure to a number of chemicals (e.g., gasoline). Because this unique protein from these animals does not occur to any appreciable extent in female rats or in both sexes of mice, kidney tumors are not seen in female rats or male and female mice. From these important sex and species differences, scientific groups and regulatory agencies have concluded that the male rat kidney tumors are not relevant to humans, a species that is also deficient in α-2microglobulin.

Certain animal strains are uniquely sensitive to certain types of cancer. For example, a large proportion of B6C3F1 mice develop liver tumors before they die, and this sensitivity appears to be due in part to the fact that the H-*ras* oncogene in this mouse strain is hypomethylated, allowing this oncogene to be expressed more easily, especially during recurrent hepatocellular injury. Similarly, 100% of strain A mice typically develop lung tumors before these animals die, and so a chemical that promotes the early development of lung tumors in this strain of mice may not produce any lung tumors in other strains that have lower lung cancer rates. The Fischer 344 rat, the rat strain commonly used in chronic cancer bioassay testing, have higher background tumors rates in certain tissues, and this difference in background tumors rates in specific organs differs from those organs with the highest background tumor rates in the B6C3F1 mouse strain, the mouse strain most commonly used in cancer bioassays. Thus, the target organs frequently sensitive to the carcinogenic effects of a particular chemical differ between the rat and the mouse. So, when tumors are induced only in those organs with high spontaneous background rates in these two rodent test species, the uncertainty regarding the human relevance of the observation increases because humans have a

TABLE 1.8 Target Organ Comparison of Tumors Induced in Chronic Cancer Bioassays

Correlations in Site-Specific Carcinogenicity[a]		
Site of Cancer	Rat Response Predicts Mouse Response	Mouse Response Predicts Rat Response
Liver	75% (25/33)	32% (25/78)
Lung	29% (2/7)	11% (2/18)
Hematopoietic system	21% (3/14)	27% (3/11)
Kidney (tubular cells)	14% (3/21)	75% (3/4)
Mammary gland	22% (4/18)	57% (4/7)
Forestomach	57% (8/14)	53% (8/15)
Thyroid gland	44% (7/16)	78% (7/9)
Zymbal gland	17% (2/12)	100% (2/2)
Urinary bladder	17% (2/12)	67% (2/3)
Skin	27% (3/11)	100% (3/3)
Circulatory system	50% (2/4)	20% (2/10)
Overall totals (all data[a])	35% (61/173)	37% (61/167)

Source: Adapted from Haseman and Lockhart (1993) and based on an examination of the results provided in 379 cancer bioassays performed by the NTP.

[a]Proportion of chemicals carcinogenic in the first species that are also carcinogenic in the second species. (Example: of the 33 chemicals inducing liver cancer in rats, 25 of these also induced liver cancer in the mouse; in contrast, of the 78 chemicals that induced liver cancer in the mouse only 25 of these also produced liver cancer in the rat).

third pattern of organ tumor incidence rates that differ from these two rodent species. The concern for human relevance may be further heightened where humans will typically be exposed at doses that are orders of magnitude lower than those required to induce tumors in either rodent species.

An illustration showing that mice and rats typically respond differently in chronic cancer bioassays is provided in Table 1.8. This table contains most of the organ comparisons made for 379 animal cancer studies undertaken by the National Toxicology Program (NTP). It shows how consistently a chemical that induced cancer in an organ of either the mouse or the rat produced the same response in the other rodent species. There are a few species-specific target organ findings where the response in one rodent species does reliably indicate where cancer might occur in the other rodent species (e.g., chemicals that produce liver cancer in rats also produce liver cancer in mice 75% of the time); however, overall there is limited predictability as to which organ might develop cancer in a particular rodent species based on the test results provided by another rodent species.

Because species differences do exist, part of any hazard or dose assessment might also be a consideration of how reproducible a specific animal response is when other studies are performed to confirm a specific toxicity. As one example of this potential problem, a group of scientists compared the results of cancer bioassays generated in the NTP to that of chemicals evaluated in other protocols as reported in the Carcinogenic Potency Database (CPDB). This comparison allowed for a determination of the percentage of replicate responses seen when a chemical was tested again for a chronic hazard of considerable public interest, cancer. Some 121 chemicals were identified that had results in both databases. Of these 121 comparisons, only 69 (57%) had concordant conclusions (i.e., the chemical had the same classification as being either a carcinogen or a noncarcinogen in the replicate test). Thus, 43% of the time 52 chemicals had discordant classifications between the two experiments. The mouse test results proved to be the least consistent between the two species with only 49% of the 70 mouse experiments showing concordant results. Of 71 rat experiments, the results were concordant 62% of the time. When evaluated by sex and species, the concordance was 46% for male mice, 36% for female mice, 55% for male rats, and 69% for female rats. Because the test comparisons involved strain differences within a species, the results were also broken down into tests using the same rat (Fischer-344) and mouse (B6C3F1) strains. This comparison resulted in the test concordance being slightly lower, with 57% concordant results for male rats, 64% for female rats, 39% for male mice, and 33% for female mice. So, the poor reproducibility of carcinogenic responses seen in the overall analysis was not caused by strain differences and variations in the rat or mouse strain being tested for a particular chemical. Less than a 100% concordance for repeated testing is not limited to cancer and may be seen for other types of toxicity tests. Failure to confirm a specific result may stem from strain, species, or test protocol differences between two or more tests of the same end point. The response differences across test species increases the uncertainty as to which result provides the most reliable reflection of the actual human response.

A similar analysis of concordant responses compared test animal responses used to predict drug safety to the human toxicities ultimately seen later during early clinical trials. This study was limited to those drug-induced toxicities that were severe enough to either terminate the development of the drug or limit the dosage used, to restrict drug use to required monitoring or restrict the targeted population. In this manner, the confusing complication of addressing the myriad of minor side effects typically associated with almost all drugs was avoided. Still, 221 examples of human toxicity for 150 different drugs were ultimately available for analysis using these selection criteria. The toxicity correlations between the human response and the animal test species were loosely defined as any effect that involved the same target organ, a choice that essentially inflates the true concordance. Still, the overall true positive concordance was stated to be only 70% when one or more species could be compared to the human response (i.e., did any test species identify the correct target organ). However, when concordance was broken down by the specific species being tested, it was found that nonrodent species had a higher concordance of 63% (primarily the dog); rodent species were concordant only 43% of the time

(here concordance was primarily from the rat). Only 36% of the time was the human toxicity concordant with both a rodent and a nonrodent species. The total animal concordance was highest among Phase 1 clinical trials (75%), but much lower in Phase 2 (58%) and Phase 3 (52%) clinical trials. Analyzing the false-negative animal findings, the authors found that when an adverse human response was not predicted, it was not the result of an inadequately low dose being tested in the animal species. Some 91% of the rodent and 90% of the nonrodent toxicology tests were judged to have been performed at a dose approximating the MTD for that species. Similarly, the animal metabolism profile correlated with that of man in 86% of the false-negative animal responses that were analyzed. Therefore, differences in metabolism were not a likely explanation for the key species-specific response differences. In fact, 89% of the time the animal and human metabolites formed are similar in both concordant and nonconcordant tests. It should be noted that the animal species that are the mainstay of toxicity studies performed on industrially and environmentally important chemicals are primarily the rat and the mouse rodent species. But these two test species resulted in poor predictions of human toxicity even though concordance was exaggerated by defining it as any effect in the same target organ, and not defined as inducing the same organ toxicity or end point.

Given these findings, it should be clear to the reader why mechanistic studies are an important component of the hazard and dose–response assessments performed in toxicology. By investigating the physiological, biochemical, and molecular changes induced in a *responsive* species/strain (one that develops the toxicity) and a *nonresponsive* species/strain (one that does not develop the toxicity), key findings may be developed that help us understand the basis for the observed species/strain differences we observe during animal testing. When this same mechanistic data is generated in human studies, the toxicologist has a better basis for predicting (or later explaining) which animal response is more likely to be relevant to humans, and thereby this leads to better species extrapolations when attempting to predict the risk or safety of human exposure to a specific chemical. Examples of this are some studies by Green examining the lung tumors induced in mice or rats by TCE. These studies provided a mechanism of action involving key biochemical and cellular responses that predicted the rank order of the tumorigenic risk would be mouse > rat >> humans. Given that the rat was the nonresponsive species even at high exposure levels, the proposed mechanism suggested TCE would not induce lung cancer even in high-dose TCE-exposed workers, a conclusion that was consistent with the available epidemiology studies.

To summarize, there are a number of important genetically driven species differences that may cause changes in the:

1. Basal metabolic rate of the test species
2. Anatomy and organ structure of the test species
3. Physiology of the test species
4. Cellular biochemistry of the test species
5. Metabolism, bioactivation, and detoxification of the chemical (see Chapter 2)
6. Toxicokinetics of the chemical (see Chapter 3)
7. One or more of these species differences may ultimately produce cellular, tissue, or organ response differences between different test species or between a test species and man

Because these differences can produce significant differences in the potency for an effect and/or in the pattern of adverse effects seen across the animal species tested, they add uncertainty to the hazard and dose–response assessment processes. Selection of the right animal to study requires a prior knowledge of the fate and effects of the chemical in humans (the goal of the animal testing), as well as its fate and effects in various animals. If data generated in only one or few test species is available, there is always uncertainty about which data will most accurately predict the human response; and there are numerous examples where either the dose–response curve or the effect are exaggerated, understated, or completely missed by the results produced in a test animal species. Determining or choosing which species best represents the human response has a great impact on the perceived and estimated risk or safety of any human exposure guideline developed from animal data. While such extrapolations may be improved where mode-of-action or mechanism-of-action data are available, or by developing information from those high-dose human exposure situations, many times these kinds of additional information simply do not exist.

1.7 ADDITIONAL FACTORS INFLUENCING HAZARD IDENTIFICATION AND DOSE–RESPONSE DATA

Route of Exposure

The exposure pathway by which a substance comes in contact with the body determines how much of it enters (rate and extent of absorption) and which organs are initially exposed to the largest concentration of the substance. For example, the water and lipid solubility characteristics of a chemical affect its absorption across the lungs (after inhalation), the skin (after dermal application), or the gastrointestinal (GI) tract (after oral ingestion), and the effect differs for each organ. The rate and site of absorption (organ) may also in turn determine the rate of metabolism and excretion of the chemical. So, changing the route of exposure may alter the dose required to produce toxicity. It may also alter the organ toxicity that is observed. For example, the organ with generally the greatest capacity for the metabolism and breakdown of chemicals is the liver. Therefore, a chemical

may be more or less toxic per unit of dosage when the chemical is given orally or peritoneally, routes of administration that ensure the chemical absorbed into the bloodstream passes through the liver before it perfuses other organs within the animal. If the capacity of the liver to metabolize the chemical within the bloodstream is great, this leads to what is referred to as a *first-pass effect*, in which the liver metabolizes a large proportion of the chemical as it is absorbed and before it can be distributed to other tissues. If the metabolism of this chemical is strictly a detoxification process, then the toxic potency of the chemical (i.e., toxicity observed per unit of dose administered) may be reduced relative to its potency when administered by other routes (e.g., intravenously). On the other hand, if the metabolism of that dose generates toxic, reactive metabolites, then a greater toxic potency may be observed when the chemical is given orally relative to inhalation, dermal, or intramuscular administrations of the chemical. (See also discussion in Chapters 2 and 3.)

As an illustration that the route of exposure may or may not affect the toxic potency of the chemical, Table 1.9 lists LD_{50} data for various routes of exposure for three different chemicals. All of these chemicals were administered to the same test species so that differences relating to the route of exposure may be compared. As this table shows, in some instances the potency changes very little with a change in the route of administration (e.g., potency is similar for the pesticide DFP: for all routes except dermal); in other instances—DDT, for example—the potency decreases 10-fold when changing the route of administration from intravenous to oral, and another 10-fold when moving from oral to dermal.

Sex

Gender characteristics may affect the toxicity of some substances. Women have a larger percentage of fat in their total body weight than men, and women also have different susceptibilities to reproduction system disorders and teratogenic effects. Some cancers and disease states are sex-linked. Large sex-linked differences are also present in animal data. One well-known pathway for sex-related differences occurs in rodents where the male animals of many rodent strains have a significantly greater capacity for the liver metabolism and breakdown of chemicals. This greater capacity for oxidative metabolism can cause the male animals of certain rodent strains to be more or less susceptible to toxicity from a chemical depending on whether oxidative metabolism represents a bioactivation or detoxification pathway for a chemical at the dose it is administered. For example, in the rat, strychnine is less toxic to male rats when administered orally because their greater liver metabolism allows them to break down and clear more of this poison before it reaches the systemic circulation. This allows them to survive a dose that is lethal to their female counterparts. Alternatively, this greater capacity for oxidative metabolism renders male rodents more susceptible to the liver toxicity and carcinogenicity of a number of chemicals that are bioactivated to a toxic, reactive intermediate during oxidative metabolism.

Age

Older people have differences in their musculature and metabolism, which change the disposition of chemicals within the body and therefore the levels required to induce toxicity. At the other end of the spectrum, children have higher respiration rates and different organ susceptibilities (generally they are less sensitive to CNS stimulants and more sensitive to CNS depressants), differences in the metabolism and elimination of chemicals, and many other biological characteristics that distinguish them from adults in the consideration of risks or chemical hazards. For example, the acute LD_{50} dose of chloroform is 446 mg/kg in 14-day-old Sprague-Dawley rats, but this dose increases to 1188 mg/kg in the adult animal.

TABLE 1.9 Effect of Route of Administration on Response (LD_{50})

Route of Administration	Methadone[a]	Strychnine[a]	DDT[a]	DFP[b]
Oral	90	16.2	420	4
Subcutaneous	48	3	1500	1
Intramuscular	—	4	—	0.75
Intraperitoneal	33	1.4	100	1
Intravenous	10	1.1	40	0.3
Intraocular	—		—	1.15
Dermal	—		3000	117

Source: Adapted from *Handbook of Toxicology*, 1956, Vol. 1.
All doses are in units of mg/kg.
[a]Rat.
[b]Rabbit.

Effects of Chemical Interaction (Synergism, Potentiation, and Antagonism)

Mixtures represent a challenge because the response of one chemical might be altered by the presence of another chemical in the mixture. A synergistic reaction between two chemicals occurs when both chemicals produce the toxicity of interest, and when combined, the presence of both chemicals causes a greater-than-additive effect in the anticipated response. Potentiation describes that situation when a chemical that does not produce a specific toxicity, nevertheless, increases the toxicity caused by another chemical when both are present. Antagonists are chemicals that diminish another chemical's measured effect. Figure 1.11 provides simple mathematical illustrations of how the effect of one or two chemicals changes if their combination causes synergism, potentiation, additivity, or antagonism, and gives a well-known example of a chemical combination that produces each type of interaction.

Modes of Chemical Interaction

Chemical interactions can be increased or decreased in one of four ways:

1. *Functional*—both chemicals affect the same physiological function.
2. *Chemical*—a chemical interaction between the two compounds affects the toxicity of one of the chemicals.
3. *Dispositional*—the absorption, metabolism, distribution, or excretion of one of the chemicals is altered by the second chemical.
4. *Receptor-mediated*—when two chemicals bind to the same tissue receptor, the second chemical, which differs in activity, competes for the receptor and thereby alters the effect produced by the first chemical.

To help illustrate the ways in which chemical interactions are increased (additive, potentiation, synergism) or decreased (antagonism), Table 1.10, adapted from a textbook on chemical interactions by Edward Calabrese, is provided. This table summarizes a few of the chemical interactions identified for drinking alcohol (ethanol) and other chemical agents that might be found in home or occupational environments.

Like alcohol, smoking may also alter the effects of other chemicals, and the incidence of some minor drug-induced side effects have been reported to be lower in individuals who smoke. For example, smoking seems to diminish the effectiveness of propoxyphene (Darvon) to relieve pain, and it lowers the CNS depressant effects of sedatives from the benzodiazepine and barbiturate families. Smoking also increases certain metabolic pathways in the liver and so enhances the metabolism of a number of drugs. Examples of drugs whose metabolism is increased by smoking include antipyrine, imipramine, nicotine, pentazocine, and theophylline. Table 1.11 summarizes a few of the chemical interactions that have been reported in aquatic toxicity studies.

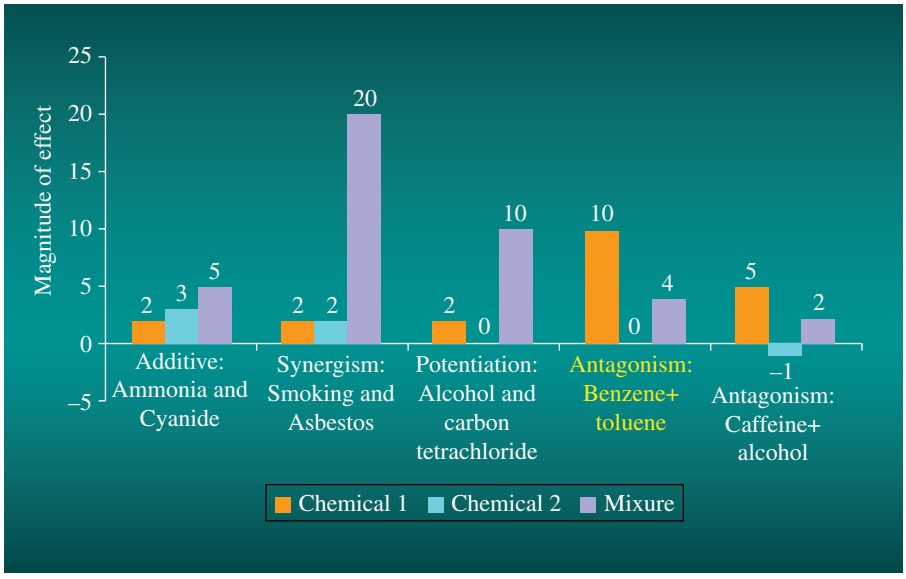


FIGURE 1.11 A simple schematic showing how a mixture formed by combining two chemicals alters the specific response of interest that is seen when the chemicals are given alone. These mixtures of chemicals the reader is likely familiar with show how the response may show additivity, synergism, potentiation, or antagonism depending on the type of interaction that occurs when the exposure contains both chemicals.

TABLE 1.10 Chemical Interactions with Ethanol

Agent	Toxic Interaction	Mode: Mechanism
Aspirin	Increased gastritis	Functional—both agents irritate the GI tract
Barbituates	Increased barbiturate toxicity	Functional/dispositional—both agents are CNS depressants; altered pharmacokinetics and pharmacodynamics of the barbiturates
Benzene	Increased benzene-induced hematotoxicity	Dispositional—enhanced benzene bioactivation to toxic metabolites
Caffeine	Caffeine antagonizes the CNS depressant effects of ethanol	Functional—both agents affect the CNS, but one is a stimulant and one is a depressant
Carbon disulfide	Enhanced CS_2 toxicity	Dispositional—increased CS_2 bioactivation and retention in critical tissues
Chloral hydrate	Increased CNS sedative effects of chloral hydrate	Functional/dispositional—both agents are CNS depressants; ethanol also alters the metabolism of chloral hydrate, leading to greater trichloroethanol accumulation
Ethylene glycol	Decreased ethylene glycol toxicity	Dispositional—ethanol inhibits the metabolism of ethylene glycol to its toxic metabolites
Nitrosamines	Increase in formation of extrahepatic tumors induced by nitrosamines	Dispositional—ethanol alters the tissue distribution of nitrosamines by inhibiting hepatic metabolism

Source: Adapted from Calabrese (1991).

TABLE 1.11 Aquatic Toxicity Interactions between Ammonia and Other Chemicals

Chemicals	Toxic End Point	Ratio of Chemical EC_{50s}	Interaction
Ammonia + cyanide	96-h LC_{50}	1:1	Additive
Ammonia + sulfide	24-h LC_{50}	1:2.2	Antagonism
Ammonia + copper	48-h LC_{50}	1:1	Additive
	48-h LC_{25}	1:1	Synergism
	48-h LC_{10}	1:1	Synergism
Ammonia + phenol	24-h LC_{50}	1:0.1	Antagonism
		1:0.7	Additive
Ammonia + phenol + zinc	48-h LC_{50}	1:1:0.5	Additive
		1:7:1	Synergism
		1:1:6	Antagonism

Source: Adapted from Calabrese (1991).

Note that when the same chemicals are present but the ratio of components present in the mixture is changed, the type of interaction observed may change. So, the interaction observed can be dose-dependent just like the toxicity is.

Genetic Makeup

We are not all born physiologically equal, and this provides both advantages and disadvantages. For example, people deficient in glucose-6-phosphate dehydrogenase (G6PD deficiency) are more susceptible than others to the hemolysis of blood by aspirin or certain antibiotics, and people who are genetically slow acetylators are more susceptible to neuropathy and hepatotoxicity from isoniazid. Table 1.12 lists some of the genetic differences that have been identified in humans and some of the agents that may trigger an abnormal response in an affected individual.

Health Status

In addition to the genetic status, the general well-being of an individual, specifically their immunological status, nutritional status, hormonal status, and the absence or presence of concurrent diseases, are features that may alter the dose–response relationship.

Chemical-Specific Factors

We have seen that a number of factors inherent in the organism may affect the predicted response; certain chemical and physical factors associated with the form of the chemical or the exposure conditions also may influence toxic potency (i.e., toxicity per unit of dose) of a chemical.

Chemical Composition The physical (particle size, liquid or solid, etc.) and chemical (volatility, solubility, etc.)

TABLE 1.12 Pharmacogenetic Differences in Humans

Condition	Enzyme Affected	Some Chemicals Provoking Abnormal Responses
Acatalasia	Catalase—red blood cells	Hydrogen peroxoide
Atypical cholinesterase	Plasma cholinesterase	Succinyl choline
Acetylation deficiency	Isoniazid acetylase	Isoniazid, sulfamethazine, procainamide, dapsone, hydralazine
Acetophenetidin-induced methemaglobinemia	Cytochrome P450	Acetophenetidin
Polymorphic hydroxylation of debrisoquine	Cytochrome P450	Encainide, metoprolol, debrisoquine, perphenazine
Polymorphic hydroxylation of mephenytoin	CYP 2C19	Mephenytoin
Glucose-6-phosphate dehydrogenase deficiency	Glucose-6-phosphate dehydrogenase	*Hemolytic anemia:* aspirin, acetanilide, aminosalicylic acid, antipyrine, aminopyrine, chloroquine, dapsone, dimercaprol, Gantrasin, methylene blue, naphthalene, nitrofurantoin, probenecid, pamaquin, primaquine, phenacetin, phenylhydrazine, potassium perchlorate, quinacrine, quinine, quinidine, sulfanilamide, sulfapyridine, sulfacetamide, trinitrotoluene

Source: Adapted from Vesell (1987).

properties of the toxic substance may affect its absorption or alter the probability of exposure. For example, the lead pigments that were used in paints decades ago were not an inhalation hazard when applied because they were encapsulated in the paints. However, as the paint aged, peeled, and chipped, the lead became a hazard when the paint chips were ingested by small children. Similarly, the hazards of certain dusts can be reduced in the workplace with the use of water to keep finely granulated solids clumped together.

Exposure Conditions The conditions under which exposure occurs may affect the applied dose of the toxicant and, as a result, the absorbed and target organ doses of the chemical. For example, chemicals bound to soils may be absorbed through the skin poorly compared to absorption when a neat solution is applied because the chemical may have affinity for, and be bound by, the organic materials in soil. The water solubility of an environmental contaminant impacts its transport through the environment and the concentrations that might be found in groundwater discharged to local waterways or used as a source of potable water. So, concentration, type of exposure (dermal, oral, inhalation, etc.), exposure pathway (soil, water, air, food, surfaces, etc.), and exposure duration (acute or chronic) are all factors associated with the exposure assessment that might alter the applied or absorbed dose of chemical.

1.8 DESCRIPTIVE TOXICOLOGY: TESTING ADVERSE EFFECTS OF CHEMICALS AND GENERATING DOSE–RESPONSE DATA

Since the dose–response relationship aids both basic tasks of toxicologists—namely, identifying the hazards associated with a toxicant and assessing the conditions of its usage—it is appropriate to summarize toxicity testing, or descriptive toxicology. While a number of tests may be used to assess toxic responses, each toxicity test rests on two assumptions:

1. *The Hazard Is Qualitatively the Same*. The effects produced by the toxicant in the laboratory test are assumed to be the same effects that the chemical will produce in humans. Therefore, the test species or organisms are useful surrogates for identifying the hazards (qualitative toxicities) in humans.
2. *The Hazard Is Quantitatively the Same*. The dose producing toxicity in animal test is assumed to be the same as the dose required to produce toxicity in humans. Therefore, animal dose–response data provide a reliable surrogate for evaluating the risks associated with different dose or exposure levels in humans.

Which tests or testing scheme to follow depends on the use of the chemical and the likelihood of human exposure. In general, part or all of the following scheme might be required in a descriptive toxicology testing program.

Level 1: Testing for acute exposure

a. Plot dose–response curves for lethality and possible organ injuries.
b. Test eyes and skin for irritation.
c. Make a first screen for mutagenic activity.

Level 2: Testing for subchronic exposure

a. Plot dose–response curves (for 90-day exposure) in two species; the test should use the expected human route of exposure.
b. Test organ toxicity; note mortality, body weight changes, hematology, and clinical chemistry; make microscopic examinations for tissue injury.
c. Conduct a second screen for mutagenic activity.

d. Test for reproductive problems and birth defects (teratology).
e. Examine the pharmacokinetics of the test species: the absorption, distribution, metabolism, and elimination of chemicals from the body.
f. Conduct behavioral tests.
g. Test for synergism, potentiation, and antagonism.

Level 3: Test for chronic exposure

a. Conduct mammalian mutagenicity tests.
b. Conduct a 2-year carcinogenesis test in rodents.
c. Examine pharmacokinetics in humans.
d. Conduct human clinical trials.
e. Compile the epidemiological data of acute and chronic exposure.

Establishing the safety and hazard of a chemical is a costly and time-consuming effort. For example, the rodent bioassay for carcinogenic potential requires 2–3 years to obtain results at a cost between $3,000,000 and $7,000,000 and when completed the results, if positive, may in the end severely limit or prohibit the use of the chemical in question. Thus, this final test may entail additional costs if now a replacement chemical must be sought that does not have significant carcinogenic activity. Figure 1.12 outlines the approximate time required to test and develop the safety of chemicals assumed to have widespread human impact.

1.9 EVIDENCE-BASED TOXICOLOGY

Toxicologists often rely heavily (and sometimes exclusively) on nonhuman studies to predict safe exposure concentrations in human populations. This type of extrapolation has many inherent issues that create uncertainty. For example, how appropriately do the *in vitro* studies reflect or predict the outcome of exposure in a whole living organism? Does the strain and/or species used in the animal toxicity studies reflect the human response qualitatively (same specific hazards and toxicities) and/or quantitatively (same dose–response data)?

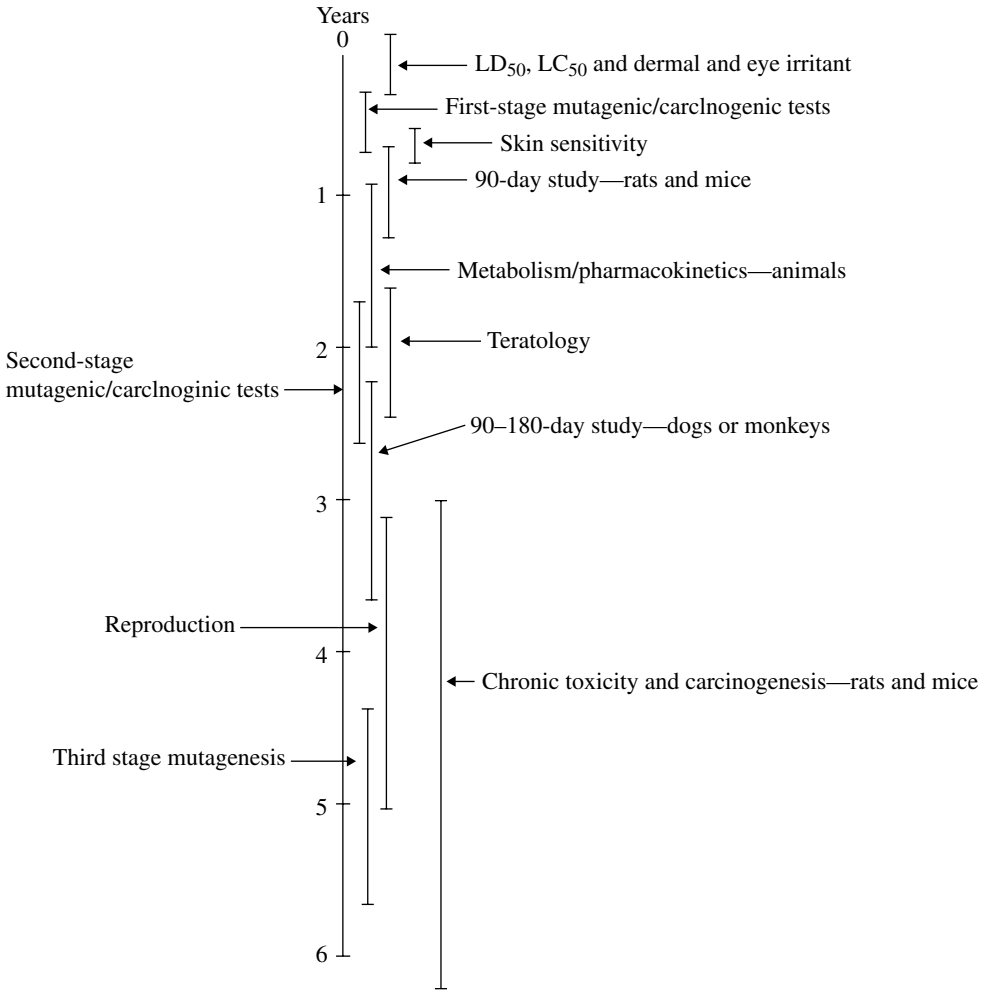


FIGURE 1.12 A timeline showing the approximate time that it might take to test a chemical having a broad exposure to the human population. The bars represent the approximate time required to complete the tests and suggest when testing might be initiated and completed.

Subsequent chapters in this book will discuss these issues in much more detail. The chapter on risk assessment will explain the deliberations required to identify the range of uncertainties with data and describe the methods used to select safety/uncertainty factors to ensure that the true human risk is not underestimated. In other chapters the reader will learn ways to limit the need for traditional animal studies and the ways that *in vitro* procedures can be performed more quickly and economically than whole animal studies. The chapters on "omics" and "computational toxicology" discusses two relatively recent subdisciplines within toxicology whose findings are driven by *in vitro* data. Other specialized disciplines within toxicology also rely heavily on *in vitro* methodologies and computer modeling. So, as animal testing becomes replaced by these approaches, the uncertainty will increase for making the correct extrapolation to human exposure.

Similar considerations occur when epidemiological data are available for a given chemical. Human exposures that occur in the workplace or community following chemical releases may afford scientists the chance to gather actual human data but this kind of information still presents challenges in extrapolation to broader populations. Epidemiology is largely an observational science and generally relies on studies where there can be limited ability to control for variables that may affect the outcome. As a result, qualitative or quantitative uncertainties may impact the use of the data in establishing acceptable human exposure values. In short, there has long been a need to better understand how to evaluate and interpret both human and animal toxicological data.

Recently two groups sought to address this issue and the term "evidence-based toxicology" (EBT) was coined to describe methods for improving the extrapolation and/or interpretation of the available epidemiology and animal toxicity data. One group has focused on improving causation analyses using human data and has proposed that toxicologists adopt the same evidence-based procedures currently used in medicine to evaluate the adverse outcomes or efficacies of different therapeutic treatments. The second group has called for the development of methods and procedures that would better characterize the utility of animal test data in a manner that ultimately could be used to improve and refine the human extrapolations/predictions from the animal data.

While this is a developing area within toxicology, it is receiving widespread interest. The initial key step in EBT is the attempt to ensure all relevant studies were considered in the hazard evaluation and subsequent risk characterization. History has shown that different "expert-based" panel evaluations for a specific chemical may reach different conclusions regarding the characterization of that chemical's hazards and possible risks. Some of these differences can arise because the data sets of studies being evaluated are not the same, and other differences ostensibly result from philosophical differences as to how specific pieces of information should be interpreted. In an attempt to reduce these problems and improve the consistency of evaluations by different scientific groups and panels, the proponents of EBT have pushed for the use of established or agreed-upon methodologies that would help improve the interpretation of test data in a manner driven more objectively by a data rating system and less by the experience and indiscernible philosophies of the panel of experts performing the evaluation. Or, stated another way, one goal of EBT is to promote a better integration of expert judgment by providing more transparent consensus methods for evaluating the evidence. In this manner differences in interpretation that arise between groups evaluating the same chemical will be more readily identified. In some instances this might help identify the key studies yet to be performed and lead to development of data critical for more accurately characterizing a chemical's hazards. For example, which mechanistic studies might be needed to understand where the threshold exists or to identify biomarkers of the disease that could be measured in an exposed population.

There are too many aspects of EBT to provide here an adequate discussion of this topic. However, the interested reader may learn more about this developing area by reading the related-articles listed in the suggested reading list at the end of this chapter.

1.10 SUMMARY

Toxicology is a scientific discipline that utilizes basic knowledge from many different but related disciplines (biology, physiology, genetics, biochemistry, etc.). The two goals or basic functions of toxicology are (1) identification of the toxicities (hazards) a chemical produces and (2) a determination of the dose range over which these hazards will be observed. Information in these two areas helps one predict what human exposures should be acceptably safe or potentially harmful, and if harmful what injuries to the exposed individual's health might be anticipated. To provide the information needed to complete these two basic functions generally necessitates the testing of animals, or the use of simpler *in vitro* tests, to predict both the hazard and dose–response outcomes in humans. The basic assumption in performing toxicological testing is that there are animal species or *in vitro* tests available that reliably and accurately provide the hazard and dose–response information we seek. Unfortunately, genetic differences across species produce differences in the anatomy, physiology, biology, and biochemistry and these species-specific differences introduce uncertainty in the premise that test animal species will respond like humans. In addition, the hazard and dose responses we observe can be changed by changes in the

testing conditions and protocol; the sex, strain, or species tested; whether exposure is to just a single chemical or that chemical as part of a mixture of chemicals or chemical exposures; the route or pathway of exposure; the genetic makeup of the human receptor exposed; and whether other chemicals the individual is also exposed to interact with and alter the toxicity of the chemical of interest. Overall, there may be considerable uncertainty in the animal-to-man extrapolation being attempted. In short, by its very nature toxicology is forced to develop data upon which uncertain extrapolations or predictions must be made. This in turn impacts how the hazards and potential risks are to be communicated to those experiencing the exposure or responsible for regulating the exposure.

SUGGESTED READING

American Conference of Government Industrial Hygienists (ACGIH). *TLV's and BEI's, Threshold Limit Values for Chemical Substances and Physical Agents*. Cincinnati (OH): ACGIH; 2012.

Ballantine B. Exposure-dose-response relationships. In: Sullivan JB, Krieger GR, editors. *Hazardous Materials Toxicology: Clinical Principles of Environmental Health*. Baltimore (MD): Williams & Wilkins; 1992. p 24–30.

Ballantine B, Sullivan JB. Basic principles of toxicology. In: Sullivan JB, Krieger GR, editors. *Hazardous Materials Toxicology: Clinical Principles of Environmental Health*. Baltimore (MD): Williams & Wilkins; 1992. p 9–23.

Ballantyne B, Marrs TC, Turner P. Fundamentals of toxicology. In: Ballantyne B, Marrs T, Turner P, editors. *General and Applied Toxicology*. New York: Stockton Press; 1993. p 3–38.

Beck BD, Calabrese EJ, Anderson PD. The use of toxicology in the regulatory process. In: Hayes AW, editor. *Principles and Methods of Toxicology*. 2nd ed. New York: Raven Press; 1989. p 1–28.

Calabrese EJ, editor. *Multiple Chemical Interactions*. Chelsea (MI): Lewis Publishers; 1991. p 467–544, 585–600.

Canadian Centre for Occupational Health and Safety (CCOHS), 2014. Available at www.ccohs.ca/oshanswers/chemicals/ld50.html. Accessed October 30, 2014.

Deschamps J, Morgan D. Information resources for toxicology. In: Ballantyne B, Marrs T, Turner P, editors. *General and Applied Toxicology*. New York: Stockton Press; 1993. p 217–230.

Eaton DL, Klassen CD. Principles of toxicology. In: Klassen CD, editor. *Casarett and Doull's Toxicology: The Basic Science of Poisons*. 5th ed. New York: McGraw-Hill; 1996. p 13–34.

Gallo MA. History and scope of toxicology. In: Klassen CD, editor. *Casarett and Doull's Toxicology: The Basic Science of Poisons*. 5th ed. New York: McGraw-Hill; 1996. p 3–12.

Garabrant DH, James RC. Trichloroethylene and cancer in humans: recognizing the need for an evidence based analysis. Toxicology 2005;212:80–84.

Guzelian PS, Victoroff MS, Halmes NC, James RC, Guzelian CP. Evidence based toxicology: a comprehensive framework for causation. Hum Exp Toxicol 2005;24:161–201.

Haseman JK, Lockhart AM. Correlations between chemically related site-specific carcinogenic effects in long-term studies in rats and mice. Environ Health Perspect 1993;101(1):50–54.

Hartung T. A toxicology for the 21st century—mapping the road ahead. Toxicol Sci 2009a;109(1):18–23.

Hartung T. Fundamentals of an evidence-based toxicology—opening statement. Hum Exp Toxicol 2009b;28:93–94.

Hartung T. Food for thought on evidenced based toxicology. ALTEX 2009c;26(1):75–82.

Hoffman S, Hartung T. Toward an evidenced based toxicology. Hum Exp Toxicol 2006;25:497–513.

James RC, Britt J, Halmes NC, Guzelian PS. Comments on recent discussions providing differing causation methodologies. Hum Exp Toxicol 2014;33:109–112.

James RC, Roberts SM, Williams PL. General principles of toxicology. In: James RC, Roberts SM, Williams PL, editors. *Principles of Toxicology Environmental and Industrial Applications*. 2nd ed. New York: John Wiley & Sons, Inc.; 2000.

Koeman JH. Toxicology: history and scope of the field. In: Niesink RJM, deVries J, Hollinger MA, editors. *Toxicology: Principles and Applications*. New York: CRC Press; 1996. p 2–15.

Loomis TA, Hayes AW. *Loomis's Essentials of Toxicology*. 4th ed. San Diego (CA): Academic Press; 1996.

Musch A. Exposure: qualitative and quantitative aspects. In: Niesink RJM, deVries J, Hollinger MA, editors. *Toxicology: Principles and Applications.* New York: CRC Press; 1996. p 16–39.

Ottobani MA. Factors that influence toxicity: how much—how often. In: Ottobani MA, editor. *The Dose Makes the Poison*. New York: Van Nostrand-Rheinhold; 1991a. p 39–54.

Ottobani MA. How chemicals cause harm. In: Ottobani MA, editor. *The Dose Makes the Poison*. New York: Van Nostrand-Rheinhold; 1991b. p 19–28.

Ottobani MA. Other factors that influence toxicity: how much—how often. In: Ottobani MA, editor. *The Dose Makes the Poison*. New York: Van Nostrand-Rheinhold; 1991c. p 55–68.

Ottobani MA. Toxicology—a brief history. In: Ottobani MA, editor. *The Dose Makes the Poison*. New York: Van Nostrand-Rheinhold; 1991d. p 29–38.

Rhodes C, Thomas M, Athis J. Principles of testing for acute effects. In: Ballantyne B, Marrs T, Turner P, editor. *General and Applied Toxicology*. New York: M. Stockton Press; 1993. p 49–88.

Sullivan JB, Krieger GR. Introduction to hazardous material toxicology. In: Sullivan JB, Krieger GR, editor. *Hazardous Materials Toxicology: Clinical Principles of Environmental Health*. Baltimore (MD): Williams & Wilkins; 1992. p 2–8.

Williams CA, Jones HD, Freeman W, Wernke MJ, Williams PL, Roberts SM, James RC. The EPC approach to estimating safety from exposure to environmental chemicals. Regul Toxicol Pharmacol 1994;20:259–280.

Williams PL. Pentachlorophenol: an assessment of the occupational hazard. Am Ind Hyg Assoc J 1982;43:799–810.

Answers to TABLE 1.3A Comparative Acutely Lethal Doses

Actual Ranking No.	LD_{50} (mg/kg)	Toxic Chemical
1	15,000	PCBs
2	10,000	Alcohol (ethanol)
3	4,000	Table salt—sodium chloride
4	1,500	Ferrous sulfate—an iron supplement
5	1,375	Malathion—a pesticide
6	900	Morphine
7	150	Phenobarbitol—a sedative
8	142	Tylenol (acetaminophen)
9	2	Strychnine—a rat poison
10	1	Nicotine
11	0.5	Curare—an arrow poison
12	0.001	2,3,7,8-TCDD (dioxin)
13	0.00001	Botulinum toxin (food poison)

Source: Adapted from Loomis and Hayes (1996).

Answers to TABLE 1.4 Occupational Exposure Limits: Aspirin and Vegetable Oil Versus Industrial Solvents

No.	Allowable Workplace Exposure Level (mg/m^3)	Chemical (Use)
1	0.05	Iodine
2	5	Aspirin (acetylsalicyclic acid)
3	10	Vegetable oil mists (cooking oil)
4	54	Trichloroethylene (solvent/degreaser)
5	55	1,1,2-Trichloroethane (solvent/degreaser)
6	75	Toluene (organic solvent)
7	147	Tetrahydrofuran (organic solvent)
8	170	Perchloroethylene (dry-cleaning fluid)
9	890	Gasoline (fuel)
10	1910	1,1,1-Trichloroethane (solvent/degreaser)

Source: American Conference of Government Industrial Hygienists (ACGIH) (2012).

2

XENOBIOTIC ABSORPTION, DISTRIBUTION, METABOLISM, AND EXCRETION

Michael R. Franklin

This chapter will familiarize the reader with:

- The nature of the biological membrane and the manner in which xenobiotics interact with and pass through it.
- The important features of the organs and the manner in which xenobiotics enter the body.
- The controlling features affecting the distribution of xenobiotics within the body.
- The major features of the enzymes catalyzing Phase I oxidation and hydrolysis and Phase II conjugation of xenobiotics, including variations present in the human population or occurring upon exposure to xenobiotics.
- The processes whereby xenobiotics and their metabolites are eliminated from the body.

This chapter examines the four cornerstones of xenobiotic pharmacokinetics—absorption, distribution, metabolism, and excretion (ADME)—on a physiological and biochemical basis. A mathematical consideration of these processes is provided in Chapter 3. Throughout the chapter there is intermittent use of words other than xenobiotic, including "chemical," "toxicant," and "drug." Where "drug" is used, there is no intent to imply that it applies only to prescribed or over-the-counter (OTC) medications; it is merely a reflection that it is in common usage in referring to all chemicals that are not part of natural mammalian biochemistry. However, much of the discussion and consideration relates to therapeutic agents, in part because generally the Food and Drug Administration (FDA) requires more information on ADME than do other regulatory agencies (e.g., Occupational Safety & Health Administration (OSHA), United States Environmental Protection Agency (USEPA)) before the chemicals they regulate that would result in human exposure to the chemical can be used. With respect to how ADME processes impact toxicities of chemicals, much of our basic understanding originated with studies examining the toxicities of chemicals ("drugs"), whether synthetic or naturally occurring in plant materials, with pharmacological activity.

The treatment of the four cornerstones varies in depth and detail, but, nevertheless, is such that all "the basics" are covered in sufficient detail that additional reading is not required to gain an overall perspective. In providing an overview, information has been gathered from innumerable sources, without providing original references for each piece of information, which would have seriously hampered the presentation and feel of the condensate. Where depth and detail are greater than might be considered warranted by an overview, they have been added to provide excitement and stimulate the curious in wanting to explore selected aspects further, and keyword searches in the electronic databases have been recommended as the place to begin. Throughout the chapter, the focus is on the human interaction with xenobiotics, and in tune with much of current research, they have expanded into aspects where the processes vary with the individual, by virtue of either genetics or concomitant or prior exposures to xenobiotics. Here in particular, the consideration is heavily weighted toward therapeutic agents because this group of xenobiotics has been the most intensively investigated but it is important to recognize that that the same principles, outcomes, and consequences apply to all classes of xenobiotics, whether present as single compounds or as mixtures. With minor variations, but with increasing differences the further removed the species is from mammals, the same principles also apply to all animal species. For nonhuman animals, and with the exception of

Principles of Toxicology: Environmental and Industrial Applications, Third Edition. Edited by Stephen M. Roberts, Robert C. James and Phillip L. Williams.

the field of veterinary medicine, the exposures are most likely to originate from the environment.

2.1 XENOBIOTIC INTERACTIONS WITH THE BIOLOGICAL (CELL) MEMBRANE

Membrane Structure

In many ways, the biological membrane can be considered a lipid bilayer (~85 Å thick) in which phospholipids are arrayed in a double back-to-back sheet with their hydrophilic (glycerol phosphate) ends oriented outward and their hydrophobic ends (fatty acids) oriented inward. While this overly simplistic rendition can be used to highlight the important features affecting the manner and rate that xenobiotics transverse membranes, in reality it is much more complicated. The lipid bilayer is not just phospholipids; cholesterol and other lipids are present and the surfaces of the sheets are coated with proteins, which may or may not be glycosylated. In addition, interspersed throughout the membrane are proteins that may be confined to one or the other side of the membrane, or may occupy both sides of the bilayer. Some of the latter may create small channels or pores in the membrane. These small (~4 Å) pores allow for the transmembrane passage of small water-soluble molecules, but they occupy only a very small fraction of the membrane surface. Therefore the majority of the membrane can be considered a lipid bilayer and it is the lipophilicity of many xenobiotic chemicals that enables them to enter or leave a cell. Entry and exit of lipophilic molecules occurs by simple diffusion (down a high-to-low chemical concentration gradient) through the central lipid bilayer. Unless aided, bulk passage of a xenobiotic chemical or toxicant is therefore a function of its relative solubility in lipid, most often expressed as its lipid/water partition coefficient. This feature is important whether considering xenobiotic entry into a cell or an organ. It applies to xenobiotics or toxicants entering the body, no matter by what "portal of entry" (i.e., organ or surface) it occurs.

Ionization of Xenobiotics

Unionized (uncharged) molecules are more lipid soluble (higher lipid/water partition coefficient) than charged molecules and so it is this form that can diffuse most easily across lipid bilayer membranes.

For chemicals with ionizable groups, the equilibrium between the unionized and ionized forms will largely determine the rate at which they can diffuse across the membrane (i.e., how much is in the unionized state). For amines, the more acidic the environment (lower pH, more H^+) the more will be the ionized form—$RNH_2 + H^+ \leftrightarrow RNH_3^+$—and therefore with less in the unionized (lipophilic) state, less will be available to diffuse through membranes. For a weak acid the more alkaline the environment (higher pH, less H^+) the more will be in the ionized form $RCOOH \leftrightarrow RCOO^- + H^+$ and with less in the unionized (lipophilic) state, less will be available to diffuse through membranes. Alternatively stated, weak acids will be most *un*ionized and diffuse most easily at low pH (acidic environment), while amines will be most *un*ionized and diffuse most easily at high pH (basic environment).

2.2 ABSORPTION

The multiple routes by which xenobiotics commonly enter or impinge upon the body and the important features of each are illustrated in Figure 2.1.

Gastrointestinal Tract

It is important to realize that compounds within the lumen of the entire gastrointestinal (GI) tract are in reality "outside" the body, that is, the cells lining the tract are an external surface and xenobiotics must pass through these cells to enter the body proper. The rate of passage is heavily dependent on xenobiotic lipophilicity at the point of contact. Entry of xenobiotics into the body via the enteral route (i.e., ingestion via the GI tract) is commonplace. In the intentional administration of chemicals with pharmacological activity ("drugs") it is considered convenient, economical, and a generally safe route. Oral administration is unsuitable for drugs that are degraded by enzymes in the digestive tract and for drugs that are not pH-stable, although various formulations and coatings can protect against exposure to stomach acid. Xenobiotic absorption from the GI tract, especially the upper regions, can be influenced by the presence of food. Food influences include physical impedance, adsorption, and complexation. Gastric emptying time and intestinal motility can influence both the rate and extent of absorption. Paradoxically, decreased gastric emptying time generally decreases the rate of absorption, a result of delayed access to the major absorptive site, that is, the small intestine. Blood circulation to the GI site of absorption affects the chemical concentration gradient that drives diffusion. Fast blood flow maintains the gradient, and slow blood flow allows the post-membrane concentration to build up and decreases the gradient.

Xenobiotics absorbed from the stomach/small intestine pass through the portal circulation directly to the liver where they may be subjected to "first-pass" metabolism prior to entry into the general circulation. First-pass metabolism serves to reduce the "bioavailability" (to the body at large) of the parent xenobiotic. For drugs, and unless the drug is bioactivated from a prodrug form, the first-pass metabolism largely accounts for any reduced oral toxicity compared with

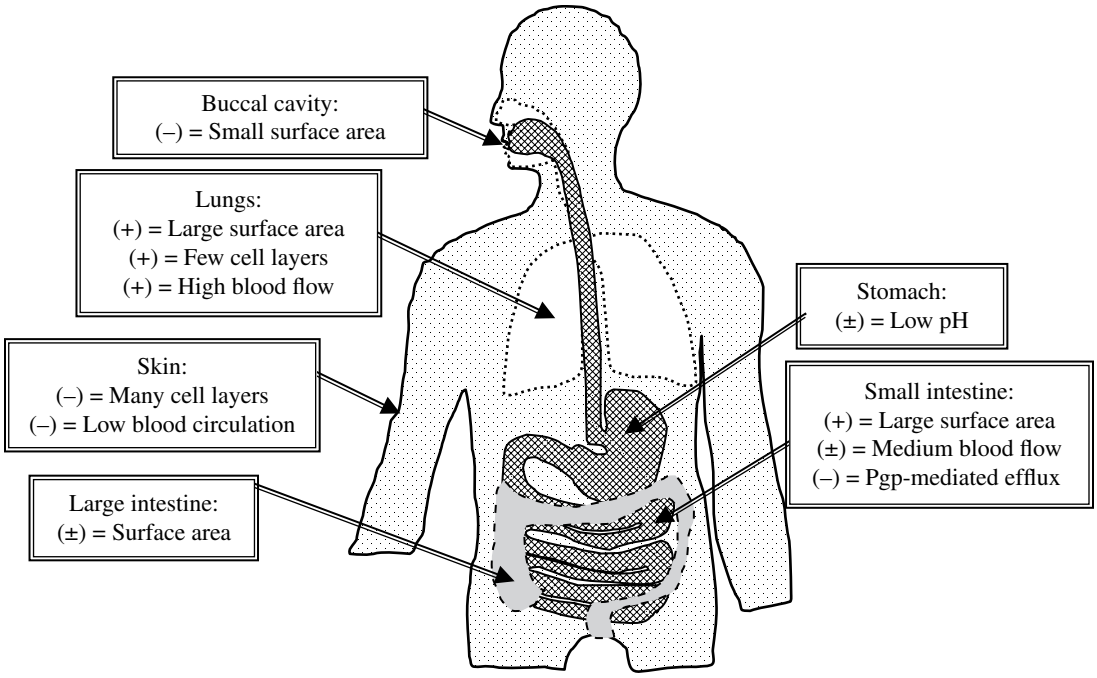


FIGURE 2.1 Key sites and affectors of xenobiotic absorption.

i.v. administration. The bioavailability to the body at large can be influenced by the portal blood flow, especially for xenobiotics with high hepatic extraction ratios. The reason why this is of greatest consequence for high extraction ratio xenobiotics is evident if one considers two xenobiotics with extraction ratios of 0.8 and 0.3. If reduced blood flow (longer hepatic "dwell" time) allows 10% greater extraction, the 20% of the 0.8 extraction ratio xenobiotic normally emerging will be reduced to 12%, an almost halving of the dose entering the general circulation. For the 0.3 extraction ratio xenobiotic, the 70% normally emerging will be reduced to 67%, a negligible change in the dose of xenobiotic entering the general circulation.

Buccal Cavity The lining of the mouth or buccal cavity is the first surface encountered by an ingested compound. The rate of diffusion is a function of, among other parameters, the surface area available, and the mouth surface area is limited. Consequently, only chemicals with a high lipid/water partition coefficient are absorbed here to any extent. Therapeutically, the absorption of nitroglycerin from a sublingual site for the quick relief of angina is evidence of this.

Stomach Most of an ingested compound quickly moves to the stomach. In the stomach, other parameters may come into play to influence the rate of absorption, most especially the change in the degree of ionization of an ionizable compound upon encountering the acid present there. It is illustrative to consider the consequences of a carboxylic acid (with a pK_a of 4.4) encountering the acid environment (pH = 1.4) of the stomach. Calculations using the Henderson–Hasselbalch equation (1.4–4.4 = log ([A−]/[HA]); −3 = log ([A−]/[HA]; 0.001 = [A−]/[HA]) show that the ionized-to-unionized ratio will be very small (1 in a 1000), so mostly the unionized compound can readily diffuse out of the stomach through the lipid bilayer of the cell membrane. However, when these compounds encounter the blood (pH = 7.4) on the other side, the degree of ionization increases dramatically. From the Henderson–Hasselbalch equation (7.4–4.4 = log ([A−]/[HA]); 3 = log ([A−]/[HA]); 1000 = [A−]/[HA]), it can be calculated that the ionized-to–unionized ratio is 1000 to 1. With the compound now mostly in the ionized form, very little will diffuse back across the membrane. The compound is "ion trapped." Overall, weak acids are well absorbed from the stomach. The reverse is true for amines; the acid environment of the stomach traps the amine within the lumen. Ion trapping occurs when a xenobiotic partitions across a membrane separating solutions having different pHs. The stomach/blood is the largest pH differential in the body and this is where ion trapping shows its greatest influence, but a blood/urine pH differential is also present in the kidney and influences elimination of acids and bases by that route. The degree of ionization is independent of either the pH or the pK_a; it is the difference between these two values that dictates this:
For acidic xenobiotics:

pH-pK_a	−3	−2	−1	0	1	2	3
% ionized	0.1	1	10	50	90	99	99.9
% unionized	99.9	99	90	50	10	1	0.1

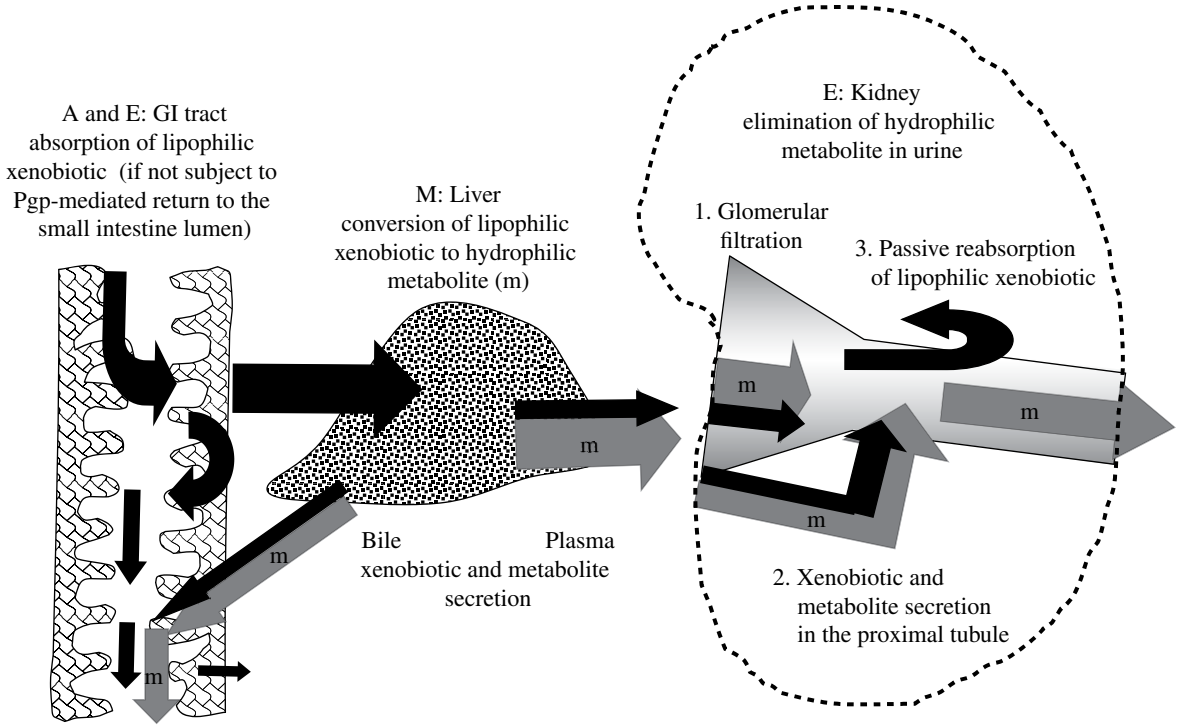


FIGURE 2.2 The interrelationship of xenobiotic absorption (A), metabolism (M), and excretion (E).

For amine xenobiotics:

pH-pK_a	−3	−2	−1	0	1	2	3
% ionized	99.9	99	90	50	10	1	0.1
% unionized	0.1	1	10	50	90	99	99.9

From the foregoing it can be easily discerned that the rate at which a xenobiotic is absorbed from the stomach is markedly affected by the pH. As such, it is also markedly affected by substances that markedly alter the pH of the stomach contents, either by direct neutralization (e.g., with sodium bicarbonate ingestion) or by reducing acid secretion (e.g., with drugs such as proton pump inhibitors).

Small Intestine For any ingested drug or other xenobiotic not absorbed from the stomach lumen, it eventually passes into the intestine where secretions raise the pH, favoring the absorption of amines over weak acids. However, weak acids are readily absorbed from the intestine. The reason for this apparent contradiction lies with another factor influencing absorption, the area of the surface through which the xenobiotic or chemical can diffuse. The surface area of the intestine is immense so even though the majority of the chemical may be in the ionized form, the molecules that are unionized (recognizing that ionization is always an equilibrium) are likely to be adjacent to a membrane through which they can diffuse. The immense surface area of the small intestine is the result of the length of the organ and the multiplier effect of the surfaces provided by the three "Fs," folds, fingers, and filaments, more properly, plicae circualares, villi, and microvilli of the brush-border enterocytes.

In a protectionist mode, the body is best served by protecting itself against the entry and circulation of xenobiotics rather than having ways and processes to remove them once inside. Within the membrane of the enterocytes are active transport "channels" that enable the cell to pump some of xenobiotics that have entered by diffusion back into the intestinal lumen (Figure 2.2). "Active transport" denotes the need for cellular energy in this process, which is provided by the hydrolysis of ATP (adenosine triphosphate). "Pump" denotes the possibility of moving the chemical up a chemical concentration gradient. (Net flow by simple diffusion is always downhill, that is, from a high concentration to a low concentration.) The channels are created by the circular arrangement of 12 or more membrane-spanning regions of a single protein termed a transporter. ATP hydrolysis at two ATP binding sites on the intracellular portion of the transporter confers the conformational change responsible for the movement of the xenobiotic through the channel. The most notable of these transporters in the intestine is *P*-glycoprotein (Pgp) or Multidrug Resistance protein 1 (MDR1). The latter designation derives from its original characterization in tumor cells, which were resistant to a variety of chemotherapeutic agents by virtue of the high expression of this efflux pump resulting in low and ineffective intracellular concentrations of the drug. MDR1 is the product of the *ABCB1*

gene and has a general preference for transporting neutral and cationic molecules. There are many ABC (ATP binding cassette) transporters, but notable in addition to MDR1 is MRP2 (MDR-Related Protein 2), which is the product of the *ABCC2* gene. This transporter has a general preference for transporting anions. The key feature of these transporters is that they enable ionized molecules to be moved (smuggled!) through a lipid bilayer and often against a chemical concentration gradient. As with enzymes, the transporters have a substrate recognition site and therefore show selectivity, may show competition and preferences between xenobiotics to be transported, and if limited in expression may show saturation with resultant zero-order kinetics (a fixed rate independent of xenobiotic concentration).

Lower GI Tract The further distal down the GI tract, the closer the lumen resembles a plain cylinder, with the folds and villi ever decreasing. Nevertheless, it does have considerable length and so continues to be a site of xenobiotic absorption. Even the rectum is utilized as an absorptive surface for the administration of some drugs, especially those likely inactivated by stomach acidity, but the absorption can be erratic and is often incomplete. Such drugs are provided in the form of suppositories and can be useful for drugs that given orally would cause severe nausea and vomiting. Substances absorbed from the lower rectum avoid the portal circulation and any liver first-pass metabolic inactivation.

Lungs

The cells lining the airway passages, like those of the skin, are external surfaces, and xenobiotics are able to pass through them to enter the body proper. These two surfaces are at opposite ends of the spectrum with regard to a parameter that affects the rate of diffusion—the distance across which a molecule diffuses. The rate of diffusion is inversely proportional to the square of the distance (double the distance, quadruple the time), and in biological terms, since most cell "lipid" membranes are of similar thickness, distance translates to the number of cell layers. In the alveoli of the lungs there are very few cell layers separating the air from the blood vessels and consequently any lipid-soluble xenobiotics in the inspired air can quickly enter the circulation. This coupled with the immense surface area presented by the alveolar structure can make the lungs an organ of risk in allowing xenobiotics into the body. The most familiar demonstration of the rapidity with which entry can occur is with gaseous or volatile liquid general anesthetics where sufficient xenobiotics can enter the bloodstream and circulate to the brain to produce anesthesia within seconds to minutes. The rapidity with which an addicted smoker can achieve sufficient central nervous system (CNS) levels of nicotine from inhaled tobacco smoke to remove any withdrawal effects is another example. The rapid absorption will of course also apply to any gases and vapors of volatile solvents. This is well known to abusers of these chemicals to obtain their quick "high" from inhalation from a closed container or bag.

Skin

For the skin, the number of cell layers between the external surface and the blood vessels is much greater than with the lung, and although lipid-soluble compounds, often organic solvents, can diffuse through, the distance to be traversed makes the diffusion rate much slower by comparison. A number of drugs have been formulated to provide delivery through transdermal patches, although the rates of absorption can be affected by regional, pathological, and individual differences in skin permeability. With injection below the epidermis, subcutaneous (SC) or intramuscular (IM), the administered drug moves by simple diffusion from the injection site into the capillaries and lymphatic system, and in both sites of administration, the blood flow to the region can markedly influence the rate of entry into the circulation. This effect is commonly seen with local anesthetics where coinjection of a vasoconstrictor serves to keep the anesthetic localized near the nerve, rather than have it washed away in the circulation.

2.3 DISTRIBUTION

Overview

Once in the bloodstream from its site of absorption or administration, several factors affect the distribution of drug or xenobiotic into the various tissue of the body. At its most basic, the distribution can be described as a volume of distribution (V_d), but this fails to identify all the contributing factors. In terms of fluid compartments into which a drug or other xenobiotic might distribute, standard values for a typical adult are plasma = 3 l, extracellular fluid = 12 l, total body water = 41 l. A compound that distributes widely therefore has a large V_d value, and if it stays within the plasma it will have a low V_d value. V_d values may also suggest other characteristics of the compound; a low V_d may indicate extensive plasma (albumin) protein binding and a high V_d may indicate extensive binding/sequestration in tissue sites.

The initial distribution to organs or tissues from absorption or administration sites is largely a function of blood flow. Highly perfused organs, that is, those receiving a high blood flow, will inevitably be exposed to high amounts of xenobiotic. How much leaves the bloodstream and enters the cells of the organ or tissue is a combination of many competing forces. For instance, if much of the compound is bound to plasma proteins or other blood components, free xenobiotic concentration driving diffusion will be low and

so will be the rate of diffusion into the tissue. The cells of many tissues have transporters embedded in their membranes that are able to return any compound that does diffuse into the cell back to the circulation. Many of these transporters are ABC transporters, closely related to the MDR1 and MRP2 described earlier that limit xenobiotic absorption from the small intestine. The brain and placenta are two organs highly protected in this manner. In the kidney, the proximal tubule cells contain on their bloodstream side transporters that serve to augment diffusion into the cell, since the kidney is the major organ of elimination of unwanted chemicals from the body. Many of these "uptake" transporters are solute carrier transporters (SLCs), which although having much structural similarity to ABC transporters, differ in their not using ATP as their energy supply. The distribution of chemicals to and through the kidney will be considered in greater detail in the section "Excretion".

In addition to the initial distribution phase, for some xenobiotics there is a subsequent redistribution. This phenomenon is seen with highly lipophilic compounds that initially enter high blood flow tissues (e.g., brain), then redistribute to tissues (e.g., muscle or fat) with lower blood flow. A drug example is thiopental, a short-acting barbiturate sedative hypnotic for which high blood flow to the brain initially delivers sufficient thiopental to the brain to cause anesthesia. Thiopental not leaving the blood for the brain on an initial pass is swept through to the body where it leaves the blood and enters into the fat. Blood now devoid of thiopental returns past the brain and reequilibration draws thiopental from the brain into the blood and this is redistributed to body fat. This continues through many circulation times until after 20–30 min insufficient thiopental remains in the brain to cause anesthesia and consciousness returns. Thus the pharmacological effect is primarily terminated not by metabolism or excretion, but by redistribution into fat; 70% of the thiopental is still present in the body 3 h after drug administration.

Albumin Binding

As mentioned earlier, the binding of drugs and other xenobiotics to blood components can be an important determinant of distribution. Binding to blood components is in the majority of cases reversible (i.e., noncovalent). Most often, binding is to the albumin component, since it is the protein of highest concentration in the plasma and a protein that also complexes with considerable amounts of lipid to which lipophilic drugs and chemicals can readily bind. The protein itself, a 66.5-kDa protein secreted by the liver, appears to have two major xenobiotic binding sites, a large flexible site that preferentially binds dicarboxylic acids and/or bulky heterocyclic molecules with a negative charge localized in the middle of the molecule and a second smaller, more narrow, and less flexible site that preferentially binds aromatic carboxylic acids with a negative charge localized at one end away from a hydrophobic center. This second site is most often associated with the binding of nonsteroidal anti-inflammatory drugs (NSAIDs). Both these sites can exhibit saturation and if multiple competing chemicals are present, displacement of one compound by another can occur. This makes prediction of the extent of albumin binding of individual components when contained in mixtures from a knowledge of binding as a single chemical extremely complex. Extensive albumin binding can restrain diffusion out of the blood into tissues. By reducing free xenobiotic concentration in the blood it can also serve to enhance absorption from, for example, the GI tract. In this consideration the binding preserves the lumen to blood high-to-low chemical concentration gradient from the "inside" by maintaining the low blood concentration low.

2.4 METABOLISM

Overview

Xenobiotic metabolism is commonly referred to as "drug" metabolism or "drug" biotransformation. It is the chemical alteration of substances by reactions in the living organism, predominantly enzyme-catalyzed. The objective of biotransformation is generally to promote the excretion of chemicals by enhancing their water solubility. Although excretion or elimination can occur without metabolism, for most compounds, especially those with considerable lipid solubility (therefore well absorbed following oral administration), it occurs subsequent to one or two types of reactions. Enhanced water solubility derives primarily from Phase II (conjugation, synthetic) reactions in which ionizable entities derived from natural body biochemicals are added to the molecule. Many conjugates (glucuronides, sulfates, amino acid conjugates) have acidic functional groups and, being extensively ionized at physiological pH, undergo less reabsorption by diffusion through lipid membranes of kidney tubule following glomerular filtration. In addition, they are subject to carrier-mediated active secretion from blood into urine (Figure 2.2). Ironically, they also require carrier-mediated efflux from the cell in which the conjugate was created. For the hepatocyte, this results in active secretion into the bile and also into the blood for subsequent renal excretion.

Many xenobiotics lack suitable functional or reactive chemical groups on which conjugation can occur and must first undergo Phase I (metabolic transformation, activation) reactions to generate them. For example, the lipid-soluble antiepileptic drug phenytoin must first be converted to 4-hydroxyphenytoin before formation of the very water-soluble 4-hydroxyphenytoin glucuronide. The same is true for benzene, which is first converted to phenol and then

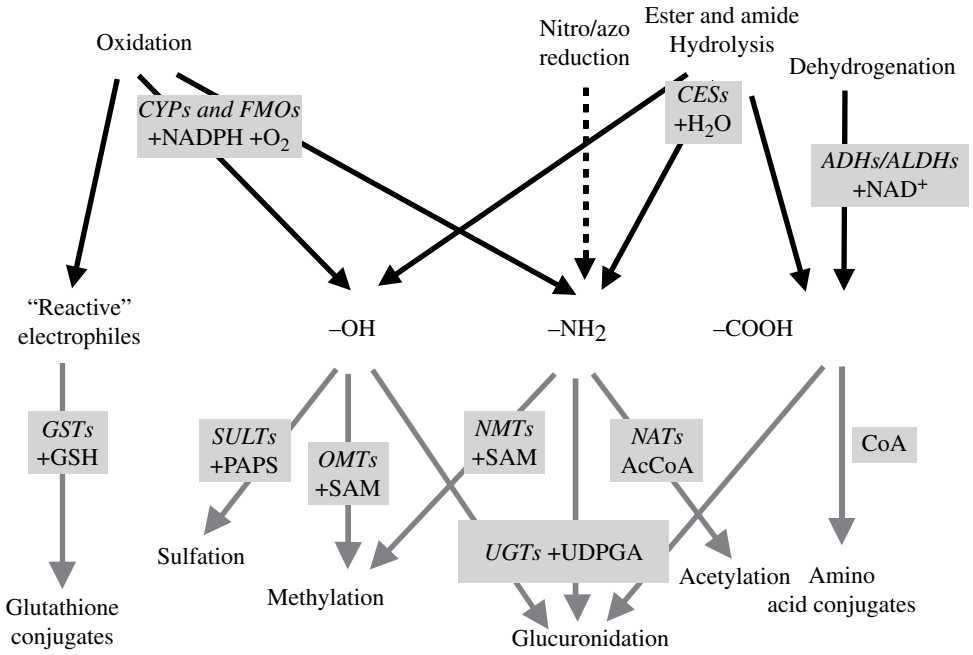


FIGURE 2.3 Major routes of xenobiotic metabolism simplified into the generation and sequestration of functional groups.

onward to phenyl glucuonide. The aqueous solubility of phenol (6.7%) far exceeds that of benzene (0.2%). In some instances, chemicals are excreted following just metabolic transformation, and small amounts of Phase I metabolites appear in urine alongside major amounts of Phase II conjugates. The two-step (Phases I and II) metabolism process largely proceeds through a limited number (three major) of functional groups (Figure 2.3) (the enzymes involved in these reactions will be considered individually later). Phase II metabolites are usually pharmacologically inactive. In addition to promoting excretion, metabolism and the increased water solubility decreases the entry of xenobiotics into cells and decreases the interaction with intracellular targets, most often receptors. Phase I metabolites may or may not be pharmacologically active. For some chemicals, reaching a state suitable for conjugation reactions requires passing through a more chemically reactive entity or intermediate (Figure 2.4). Highly reactive metabolites, most often generated from Phase I oxidations, are occasionally produced and these can cause tissue damage. Mutagenic and carcinogenic epoxides generated by P450s are good examples of such intermediates; others are shown in Table 2.1. Reactive intermediates may cause enzyme inactivation, membrane lipid peroxidations resulting in membrane alterations, and changes in DNA. Reactive intermediates have been implicated in carcinogenesis, tissue necrotic reactions, and tissue allergic responses.

Xenobiotic biotransformation takes place in almost all organs and tissues (liver, skin, GI tract, lungs, kidney, blood, etc.); however, the liver is quantitatively the most important tissue for xenobiotic metabolism especially because of its high expression levels of many xenobiotic-metabolizing enzymes. The liver with its portal blood supply from the GI tract is uniquely placed to protect the body from exogenous, possibly toxic chemicals. If the liver can rapidly metabolize the xenobiotic following absorption from the GI tract, a significant portion of the xenobiotic may be inactivated before it can produce a therapeutic or toxic effect in the remainder of the body.

Within the liver, a major subcellular site of xenobiotic biotransformation is the endoplasmic reticulum. Among the most common/important xenobiotic-metabolizing enzymes located here are cytochrome P450s and flavin monooxygenases (FMOs), carboxylesterases, and glucuronosyltransferases. When liver is homogenized and ruptured cells are differentially centrifuged, fragments of the endoplasmic reticulum with xenobiotic-metabolizing capabilities are isolated in the fraction called microsomes (artifacts of cell disruption), a term often used to identify the cellular location of xenobiotic-metabolizing enzymes. Differential centrifugation of liver tissue gently homogenized in 0.25 M sucrose (so as to preserve mitochondria intact) can yield other fractions containing xenobiotic-metabolizing enzymes; 18,000 g for 20 min will pull down intact heavy and light mitochondria into a pellet, and microsomes require 105,000 g for 60 min to sediment, leaving the cytosolic (soluble) fraction as a supernatant. Note that a 20-min 9000 g supernatant, often termed an S9 fraction, a fraction that is utilized in some xenobiotic metabolism studies and mutagenic assays ("Ames test"), contains both microsomes (endoplasmic reticulum) and cytoplasm (and some light mitochondria) and therefore enzymes contained therein.

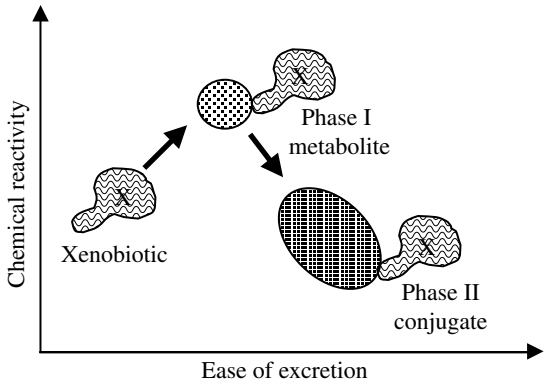


FIGURE 2.4 Changes in chemical reactivity and excretability with metabolism.

TABLE 2.1 Toxic Xenobiotics Classed by Reactive Metabolites/Intermediates

Epoxides
Aflatoxins B_1 and B_2, benzo(a)pyrene and benzo(e)pyrene, chrysene, 7,12-dimethylbenzanthracene, 3-methylcholanthrene
Quinones
Adriamycin, *o*- and *p*-benzoquinone, bleomycin, menadione, mitomycin c, 1,2-naphthoquinone, streptonigrin
Carbonium ions
2-Acetylaminofluorene (2AAF), dimethylnitrosoamine, nitrosonornicotine, procarbazine, pyrrolizidine alkaloids
Imines
Acetaminophen, amodiaquine, 2,6-dimethylaniline, ellipticine acetate, 3-methylindole, nicotine, phencyclidine
Nitrenium
2-acetylaminofluorene (2AAF), 4-aminobiphenyl, 2-aminonaphthalene, 2-aminophenanthrene, benzidine
Acyl glucuronides
Bilirubin, clofibric acid, diflunisal, indomethacin, tolmetin, valproic acid, zomepirac
Glutathione adducts
Chlorotrifluorethylene, 1,2-dibromo-3-chloropropane, dibromoethane, *N*-(3,5-dichlorophenyl)succimide, hexachlorobutadiene, tetrachloroethylene, tetrafluoroethylene, trichlorethylene, tris(2,3-dibromopropyl) phosphate

Metabolism Enzymes and Reactions

Phase I Oxidations There are two groups of enzymes in the liver that are largely responsible for the oxidation of xenobiotics, both termed monooxygenases, that are exclusively localized to the endoplasmic reticulum. One group consists of heme-containing proteins termed cytochrome P450s (abbreviated CYPs), and about 12 of these are important in xenobiotic metabolism. The second group consists of enzymes termed flavin monooxygenases (abbreviated FMOs) and, of the five forms in the human genome, FMO3 is the major form present in the liver. Both CYPs and FMOs require NADPH and O_2 for their catalytic function, but they differ in the reaction mechanism, and also in the range of xenobiotics they oxidize. In a minority of cases, a compound can be metabolized by either enzyme (e.g., nicotine, clozapine). Additional enzymes also catalyze the oxidation of xenobiotics, most notably dehydrogenases or monoamine oxidases (MAOs). These enzymes have wider distribution in subcellular compartments than CYPs and FMOs.

Flavin Monoxygenases FMOs oxidize via the formation of a hydroperoxyflavin species (NADP–FADHOOH). Substrates are compounds that contain a nucleophilic heteroatom (nitrogen, sulfur) presenting a lone pair of electrons. Nitrogen oxidation can occur with both tertiary (e.g., *N*-dimethylaniline, imipramine, amitryptyline) and secondary (e.g., *N*-methylaniline, deispramine, nortriptyline) amines. Sulfur oxidation can occur with thiols (e.g., dithiothreitol, β-mercaptoethanol), sulfides, (e.g., dimethylsulfide), thioamides (e.g., thioacetamide), and thiocarbamides (e.g., thiourea, propylthiouracil, methimazole). A notable feature of FMO3, the dominant liver form, is a polymorphism (a 551C-T mutation in exon 4) that changes a proline residue (153) to a leucine, resulting in an inactive enzyme. This polymorphism is associated with a trimethylaminuria or fish odor syndrome phenotype, since trimethylamine (from choline, lecithin, and carnitine metabolism) cannot be converted to a nonodorous metabolite.

Among pharmacological classes containing drugs metabolized in part by FMOs are psychotropic therapeutics (clozapine, chlorpromazine, fluoxetine, imipramine), H2 receptor antagonists (cimetidine, ranitidine), a thioureylene antithyroid agent (methimazole), antihistaminic agents (promethazine, brompheniramine), an antiarrhythmic agent (verapamil), an antifungal agent (ketoconazole), and a cancer chemotherapeutic (tamoxifen).

Cytochrome P450s "Cytochrome P450" is a term for a large superfamily family of related hemoproteins. The reactions they catalyze are often described by their chemistry, but with only a few exceptions, most are oxidations, as can be seen by the addition of an oxygen atom in the basic chemical reactions (Table 2.2). The xenobiotic oxidation occurs via a cyclic set of orchestrated steps. First the xenobiotic substrate binds to the P450 at a substrate recognition site, which alters the protein structure sufficiently for the heme iron atom of the enzyme to receive an electron from an FAD- and FMN-containing flavoprotein, NADPH P450 reductase (CYPOR). The reduction of the heme iron from ferric to ferrous facilitates the binding of molecular oxygen, and with an additional electron from CYPOR, and a series of electronic rearrangements, one atom of the molecular oxygen is bonded to the xenobiotic, the other oxygen atom (with protons) becoming a

TABLE 2.2 Cytochrome P450–Catalyzed Oxidation Reactions

Reaction name	Reaction Formula	Xenobiotic Example(s)
Aromatic hydroxylation	$C_6H_5X \rightarrow HOC_6H_4X$	Benzene, toluene, coumarin, naphthalene, benzopyrene, fluorene, phenytoin
(Aromatic) epoxidation	$C_6H_5X \rightarrow OC_6H_5X$	Benzopyrene, carbamazepine, coumarin, aflatoxin B1, styrene
Alicyclic oxidation	$RCH_3 \rightarrow RCH_2OH$	Toluene, phenobarbital *N,N*-diethyl-*m*-toluamide (DEET)
O-dealkylation[a]	$ROCH_3 \rightarrow ROCH_2OH \rightarrow ROH + HCHO$	Dextromethorphan
N-dealkylation[a]	$RNHCH_3 \rightarrow RNHCH_2OH \rightarrow RNH_2 + HCHO$	Diazepam, DEET caffeine, dextromethorphan,
Oxidative deamination	$R_2CHNH_2 \rightarrow R_2C(OH)NH_2 \rightarrow R_2CO + NH_3$	Dextro-amphetamine
Sulfoxidation	$RSR \rightarrow RS(OH)R \rightarrow RSOR + H+$	Chlorpromazine
Desulfuration	$R_2CS \rightarrow R_2CO + S$	Parathion, malathion, thiopental
N-oxidation	$R_3N \rightarrow R_3NOH \rightarrow R_3NO + H+$	Guanethidine, 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone, (NNK).

[a]Alkyl group removed during dealkylations is most often a methyl or an ethyl.

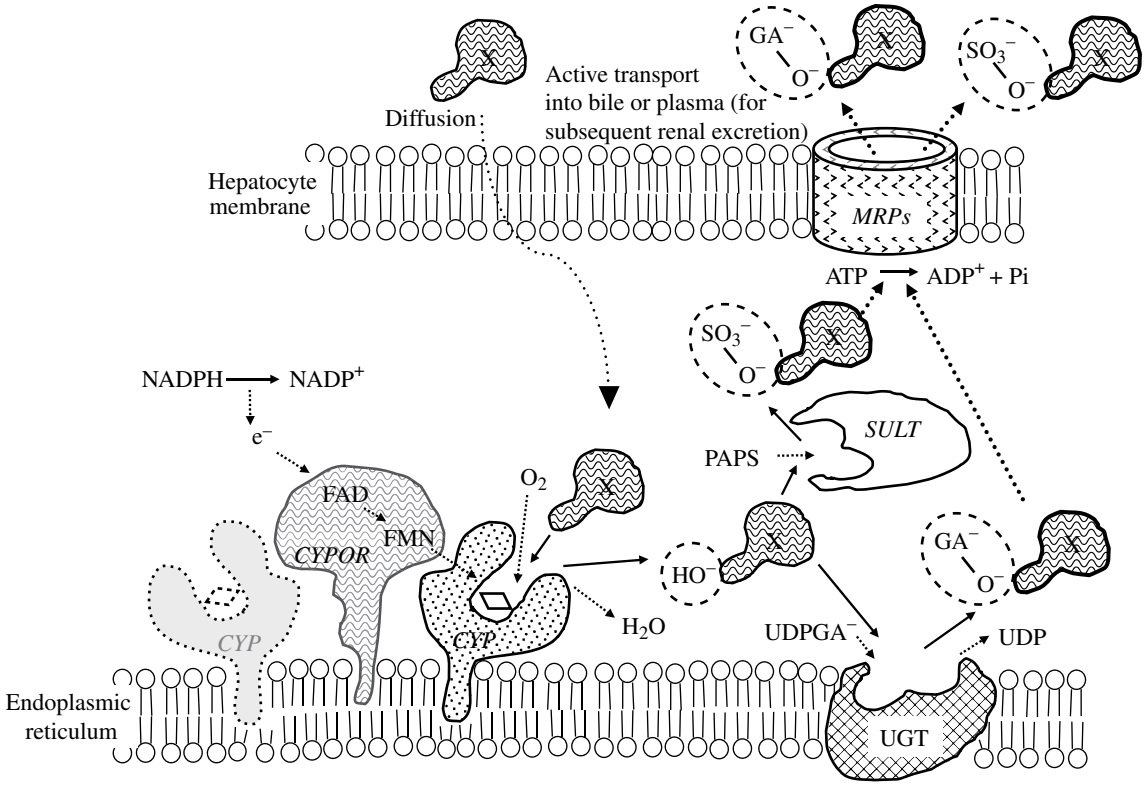


FIGURE 2.5 Hepatocyte xenobiotic metabolism and excretion.

molecule of water. The components of this oxidation cycle are illustrated in Figure 2.5.

Cytochrome P450 family members can differ markedly in their substrate and inhibitor specificity/selectivity, abundance, transcription control, and polymorphism frequency. In any extended discussion of this most important group of xenobiotic-metabolizing enzymes, it is therefore necessary to refer to them individually. The developed nomenclature, based on amino acid homology, is as follows:

1. The family, defined as >40% homology between members, is given a number, CYP1, 2, 3, and so on, for example, CYP2.

TABLE 2.3 Notable Xenobiotic Substrates of Cytochrome P450s

CYP	
1A2	Amitriptyline, benzo[a]pyrene, caffeine, clomipramine, clozapine, cyclobenzaprine, estradiol, fluvoxamine, haloperidol, imipramine (N-demethylation), mexiletine, naproxen, olanzapine, ondansetron, phenacetin (O-demethylation to acetaminophen), propranolol, riluzole, ropivacaine, tacrine, theophylline, tizanidine, verapamil, (*R*)warfarin, zileuton, zolmitriptan.
2A6	Nicotine
2B6	Bupropion, cyclophosphamide, efavirenz, ifosfamide, methadone.
2C8	Paclitaxel, torsemide, amodiaquine, cerivastatin, repaglinide.
2C9	Amitriptyline, celecoxib, diclofenac, fluoxetine, fluvastatin, glibenclamide, glimepiride, glipizide, glyburide, ibuprofen, irbesartan, lornoxicam, losartan, meloxicam, *S*-naproxen, nateglinide, phenytoin (4-hydroxylation), rosiglitazone, suprofen, tamoxifen, tolbutamide, torsemide, *S*-warfarin
2C19	Amitriptyline, carisoprodol, citalopram, chloramphenicol, clomipramine, cyclophosphamide, diazepam, hexobarbital, imipramine (N-demethylation), indomethacin, lansoprazole, *S*-mephenytoin, mephobarbital, moclobemide, nelfinavir, nilutamide, omeprazole, pantoprazole, phenobarbital, phenytoin, primidone, progesterone, proguanil, propranolol, rabeprazole teniposide, *R*-warfarin (8-hydroxylation)
2D6	Amitriptyline, amphetamine, aripiprazole, atomoxetine, bufuralol, carvedilol, chlorpheniramine, chlorpromazine, clomipramine, codeine (O-demethylation), debrisoquine, desipramine, dextromethorphan, duloxetine, flecainide, fluoxetine, fluvoxamine, haloperidol, imipramine, lidocaine, metoclopramide, *S*-metoprolol, mexiletine, nebivolol, nortriptyline, ondansetron, oxycodone, paroxetine, perphenazine, phenacetin, promethazine, propafenone, propranolol, risperidone (9-hydroxylation), sparteine, tamoxifen, thioridazine, timolol, tramadol, venlafaxine, zuclopenthixol
2E1	Acetaminophen (to NAPQI), aniline, benzene, chlorzoxazone, *N,N*-dimethyl formamide, enflurane, ethanol, halothane, isoflurane, methoxyflurane, sevoflurane, theophylline (8-hydroxylation)
3As	Alfentanil, L-alphacetylmethadol, alprazolam, amlodipine, aprepitant, aripiprazole, astemizole, atorvastatin, buspirone, cafergot, caffeine, cerivastatin, chlorpheniramine, cilostazol, clarithromycin, cocaine, codeine (N-demethylation), cyclosporine, dapsone, dexamethasone, dextromethorphan, diazepam (3-hydroxylation), diltiazem, docetaxel, domperidone, eplerenone, erythromycin, estradiol, felodipine, fentanyl, finasteride, haloperidol, hydrocortisone, imatinib, indinavir, irinotecan, lidocaine, lercanidipine, lovastatin, methadone, midazolam, nateglinide, nelfinavir, nifedipine, nisoldipine, nitrendipine, ondansetron, pimozide, progesterone, propranolol, quetiapine, quinidine (3-hydroxylation), quinine, risperidone, ritonavir, salmeterol, saquinavir, sildenafil, simvastatin, sirolimus, tacrolimus (FK506), tamoxifen, paclitaxel, terfenadine, testosterone, trazodone, triazolam, verapamil, vincristine, zaleplon, ziprasidone, zolpidem

2. The subfamily with >59% homology between members is given a letter, CYPn A, B, C, and so on, for example, CYP2D.
3. The individual identifier based on the order of discovery independent of animal species is given a second number, CYP(n,lttr) 1, 2, 3, and so on, for example, CYP2D6. Thus humans do not have six forms of CYP2D; CYP2D1–5 occur in other animal species
4. Any polymorphism is identified with an asterisk and a number, CYP(n,lttr,n)*1,2,3, and so on, for example, CYP2D6*4.

The normal abundance of the cytochrome P450s of greatest consequence for the oxidative metabolism of xenobiotics in the liver and the extent to which they participate in the metabolism of drugs that undergo P450 catalyzed oxidation varies widely. CYP3A4 and CYP3A5 are the most abundant (about 40% of the total) and metabolize about 40% of the therapeutic agents that undergo oxidation. Members of the CYP2C subfamily (CYP2C8, 9, and 19) collectively are about 25% of the total and also metabolize about 25% of the drugs that undergo oxidization. The respective approximate numbers are 17 and 10% for CYP1A2, 6 and 3% for CYP2A6, 2 and 2% for CYP2B6, and 10 and 5% for CYP2E1. The biggest imbalance is seen with CYP2D6, which although present at less than 5% of the total, is responsible for the metabolism of about 20% of the drugs that undergo CYP-dependent oxidative metabolism.

Each cytochrome P450 shows some degree of substrate selectivity (Table 2.3). From Table 2.3, it is evident that CYP enzymes also show some promiscuity in their substrate selectivity. Select examples among drugs include amitriptyline (metabolized by CYP1A2, CYP2Cs, and CYP2D6), cyclophosphamide (metabolized by 2B6 and 3A4), haloperidol (metabolized by 2D6 and 3A4), imipramine (metabolized by 2C19 and 2D6), mexiletine (metabolized by 1A2 and 2D6), and phenytoin (metabolized by 2C9 and 2C19). Even for a single substrate, a single P450 can oxidize at different sites on the molecule if the molecule is able to orient in various positions within the active site. The coverage that

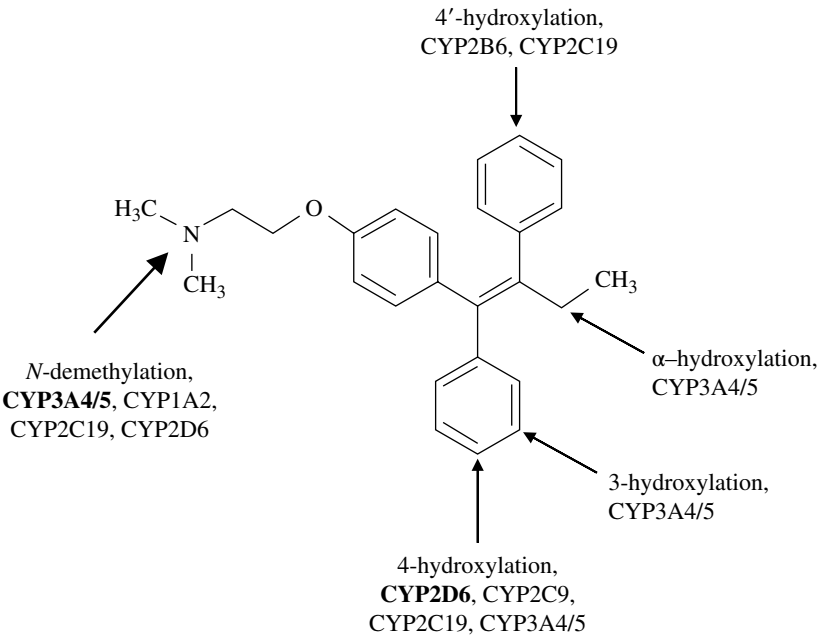


FIGURE 2.6 Multiple-site and multiple-enzyme participation in metabolism.

the multisite, multienzyme metabolism provides can be readily illustrated with the metabolism of the breast cancer chemotherapeutic drug, tamoxifen (Figure 2.6).

Polymorphisms The coverage for metabolizing tamoxifen notwithstanding, enzyme polymorphisms for substrates with lesser coverage can substantially alter pharmacokinetics. Polymorphisms are important in understanding some of the variations seen in CYP-dependent metabolism. This understanding is especially important when polymorphisms result in decreased rates and/or extent of metabolism. Several such examples are known for a polymorphism in CYP3A5. CYP3A5*3 is a low- (no-) activity enzyme due to an A/G nucleotide change (from wild-type CYP3A5*1) in intron 3 of the gene that creates a splice variant. A significant difference (increase) was observed in the blood levels of cyclosporine, an immunosuppressive drug, in the early period after renal transplantation between patients with the CYP3A5*1/*1 genotype and those carrying the CYP3A5*3 variant allele. Concern in this example arises because excessive levels of cyclosporine result in nephrotoxicity, renal vascular damage, and hypertension. A typical effect of an activity-reducing polymorphism on pharmacokinetics is shown in Figure 2.7 Simvastatin, a 3-hydroxy-3-methylglutaryl–coenzyme A (HMG-CoA) reductase inhibitor, with beneficial effects on coronary disease and mortality rates in patients with hypercholesterolemia, undergoes extensive first-pass metabolism (primarily mediated by CYP3A4/5) to inactive metabolites in the intestinal wall and liver. The mean area under the plasma concentration–time curve of simvastatin in the CYP3A5*3/*3 carriers is significantly

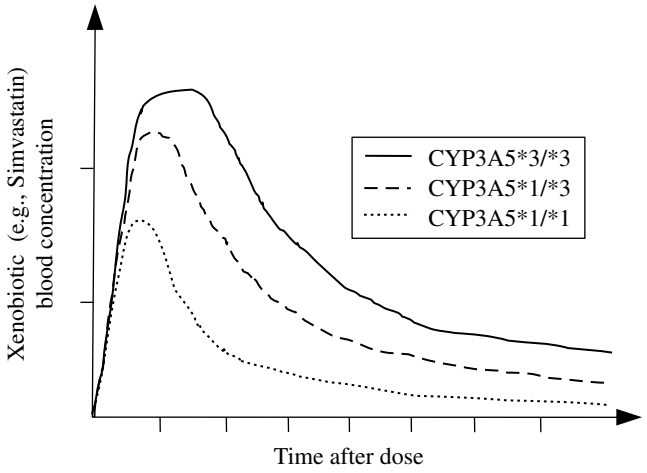


FIGURE 2.7 Activity-decreasing polymorphisms and pharmacokinetics.

higher than in the CYP3A5*1/*1 carriers, with heterozygote CYP3A5*1/*3 carriers as the intermediate.

Polymorphisms have been found for most if not all cytochrome P450s. "Functional" effects and the frequency of the polymorphisms for the major cytochrome P450s varies even among closely related members. CYP3A5 polymorphisms occur with high frequency and have significant functional outcomes whereas CYP3A4 polymorphisms occur with low frequency and appear to have little functional outcome. In the CYP2 family, CYP2D6 and CYP2C19 polymorphisms occur with high frequency and have significant functional outcomes. While also having significant functional

outcomes, CYP2C9 polymorphisms have a much lower frequency. The significant functional outcomes for CYP2C9 can be seen with the widely prescribed anticoagulant warfarin (a vitamin K antagonist) where an excessive dose can cause hemorrhaging, and an inadequate dose can result in a thromboembolic event. CYP2C9 is responsible for the oxidation (aromatic hydroxylation on the chromene ring) of warfarin, and studies have shown that low-activity *3/*3 patients (2 copies of *3 variant) may require dose reductions of up to 80% of normal (wild type; *1/*1). The low activity of CYP2C9*3 arises from a DNA base coding change that results in the presence of isoleucine instead of leucine at residue 359, a relatively conservative change. Another low-activity variant (CYP2C9*2) arises from a cysteine instead of an arginine at residue 144, an amino acid switch also seen (at residue 487) with CYP2B6*5. The majority of cytochrome P450 polymorphisms are the result of amino acid differences, although it should be noted that not all amino acid differences result in major or noticeable changes in enzyme activity. For some enzymes, however, base changes in the DNA create a premature stop codon (e.g., CYP2C19*3), and in others (e.g., CYP2C19*2, CYP2D6*4, CYP3A5*3) they result in splicing defect resulting in mRNA deficits. Some polymorphisms are the result of changes in the noncoding region of the gene, which result in altered transcription, in some cases higher than normal (seen with CYP1A1 and CYP1A2) and in others lower than normal (CYP2A6*9). Consideration of these variations from "normal" is not only important for the metabolism of pharmaceutical drugs, but it can be of consequence in the metabolism of workplace chemicals and environmental contaminants, especially those that are carcinogens.

The functional outcomes of consequence for enzyme polymorphisms are seen with population studies. Such studies often uncover ethnic or geographical differences, a confounding factor in drug development and usage, as well as in the study of diseases thought to arise from metabolism of environmental chemicals and foodstuffs. With respect to the CYP2C9 polymorphisms mentioned earlier, while both the CYP2C9*2 and CYP2C9*3 alleles are common in Caucasians, CYP2C9*3 is low and CYP2C9*2 is almost absent in East Asian populations.

As a further example, CYP2D6 is a highly polymorphic enzyme with over 50 documented, although the majority (~95% in the Caucasian population) of the poor metabolizer phenotype can be attributed to less than 6 mutant alleles. The most common mutant allele is CYP2D6*4 with CYP2D6*3 and CYP2D6*5 having a very much lower frequency. When drug metabolism of a CYP2D6 substrate (groundbreaking studies utilized the antihypertensive drug debrisoquine) is monitored across a sufficiently large population, four phenotypes can be identified: poor, intermediate, normal (extensive), and ultrarapid metabolizers (Figure 2.8a). For CYP2D6-dependent metabolism, there is a higher incidence of poor metabolizers in Asians, Pacific Islanders, Africans, and African Americans. About 35% of the Caucasian population are intermediate metabolizers with one functional and one mutant *CYP2D6* gene. Ultrarapid metabolizers have duplicate or multiple copies of the entire *CYP2D6* gene. The consequences of these various CYP2D6 phenotypes on pharmacokinetic profiles is also shown (Figure 2.8b) using nortriptyline (a tricyclic antidepressant) as an example. With no functional *CYP2D6* genes, poor metabolizers may reach toxic plasma levels from a standard dose, while ultrarapid metabolizers may result in area-under-the-curve (AUC) plasma levels ineffective for the treatment of depression. CYP2D6 metabolism phenotypes are also important in whether patients derive analgesia from codeine. CYP2D6 is responsible for the conversion (O-demethylation) of codeine (a low-affinity opioid agonist) to morphine (a high affinity agonist) (Figure 2.8c). While extensive and poor metabolizers derive the same degree of pain relief from a standard dose of morphine (thus eliminating receptor [pharmacodynamic] differences), poor metabolizers frequently derive little or no benefit from codeine. (In the absence of activating CYP2D6 metabolism, codeine can be inactivated via N-demethylation by CYP3A4 to norcodeine.)

It should be noted that not all enzyme activity-decreasing polymorphisms are detrimental, at least to therapeutic outcomes. Polymorphisms in CYP2C19 result in rapid (*1/*1), intermediate (*1/*2, *1/*3), and poor metabolizers (*2/*2, *3/*3, *2/*3) of proton pump inhibitor drugs such as omeprazole (O-demethylation) and lansoprazole (hydroxylation on benzimidazole) that are used in the therapy of *Helicobacter pylori* infection and gastroesophageal reflux diseases. Poor metabolizers show a larger plasma drug AUC that in turn increases the pharmacological effect, the suppression of acid secretion. The resulting higher average pH of the stomach lumen results in a better stomach ulcer healing and eradication rate. In common with CYP2D6 and CYP2C9 polymorphisms, ethnic differences occur in the frequency of CYP2C19 polymorphisms. The CYP2C19*2 allele is the most common and tends toward higher levels in Indians and East Asian populations as compared to Caucasians. CYP2C19*3 allele is rare in Caucasians, Africans, and Indians but is quite common in East Asians.

Inhibition In addition to polymorphisms, decreased CYP metabolism often arises from "drug–drug" interactions (DDIs). If concurrently administered drugs or xenobiotics encountered in the diet and environment are metabolized predominantly by different P450s, there may be little or no inhibition-based DDIs (no interference with each other's metabolism). However, interference with metabolism is quite common and can be of considerable consequence, increasing the plasma xenobiotic AUC by more than fivefold (Table 2.4). Some xenobiotics inhibit more than one cytochrome P450, although the relative inhibitory activity toward each CYP may differ. When considering exposure to

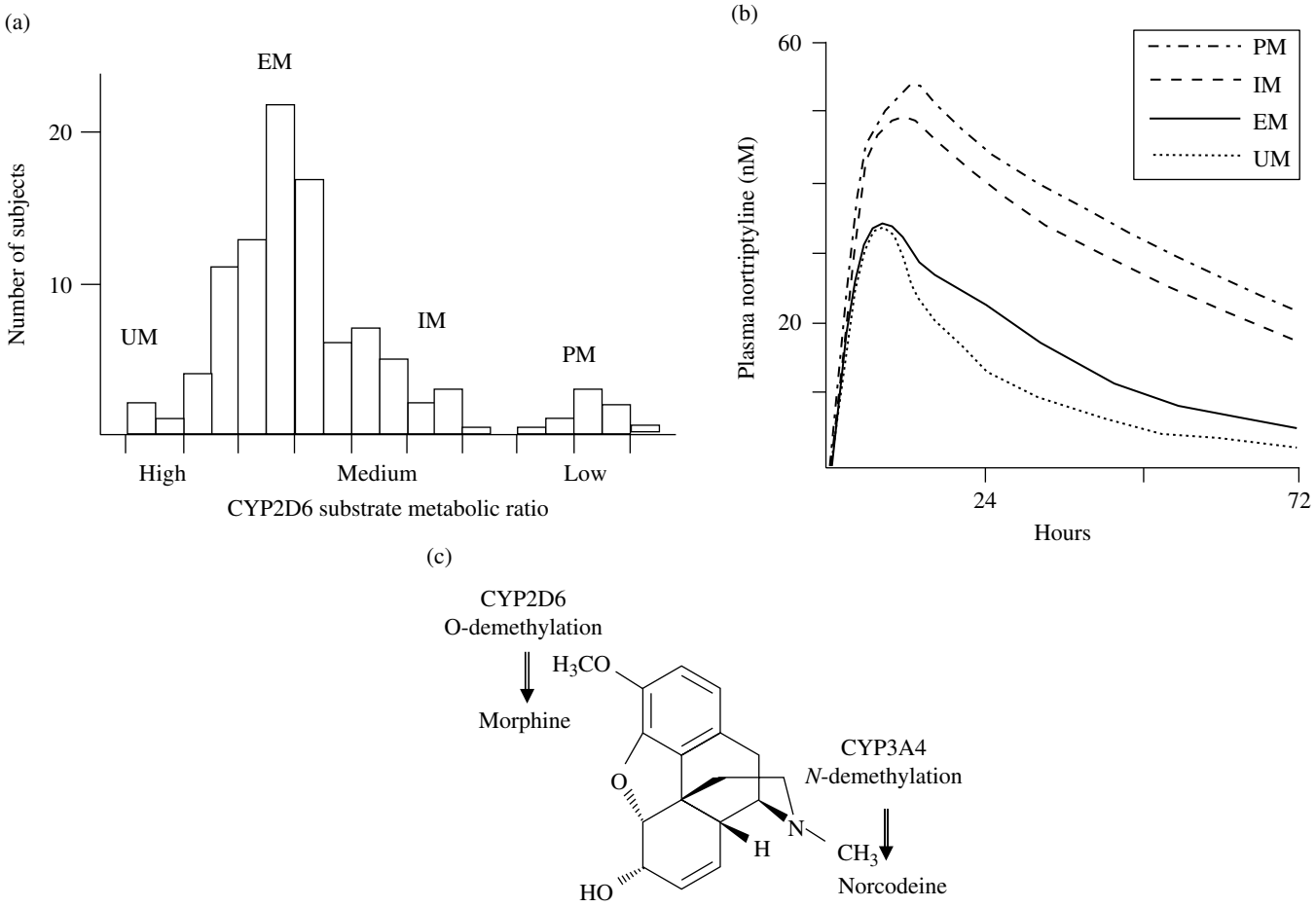


FIGURE 2.8 CYP2D6 polymorphism considerations.

TABLE 2.4 Compilation[a] of Drugs Showing Clinically-Significant Inhibitory Drug Interactions[b] At Cytochrome P450s

CYP	Therapeutic Agent[c]
1A2	*Acyclovir, cimetidine,* ciprofloxacin, *famotidine*, **fluvoxamine**, mexiletine, *norfloxacin,* propafenone, *verapamil,* zileuton
2C8	**Gemfibrozil**, *trimethoprim*
2C9	Amiodarone, fluconazole, oxandrolone, *sulfinpyrazone*
2C19	**Omeprazole**
2D6	*Amiodarone,* duloxetine, **fluoxetine, paroxetine, quinidine**, *sertraline,* terbinafine
3A	Amprenavir, aprepitant, **atazanavir, clarithromycin**, *cimetidine*, diltiazem, erythromycin, fluconazole, fosamprenavir, grapefruit juice, **indinavir, itraconazole, ketoconazole, nefazodone, nelfinavir, ritonavir, saquinavir, telithromycin**, verapamil

[a]FDA Guidance for Industry Publication (Draft Guidance), September 2006, Clinical Pharmacology.
[b]At three degrees of severity, indicated by font style: **Bold = Strong inhibition**, causing a >5-fold increase in plasma AUC values or more than 80% decrease in clearance, standard = moderate inhibition causing a >2-fold increase in the plasma AUC values or 50–80% decrease in clearance, and *italic = Weak inhibition causing a >1.25-fold but <2-fold increase in plasma AUC or 20–50% decrease in clearance.*
[c]Drugs underlined have been suggested for use in *in vivo* studies, as have fluvoxamine and moclobemide for CYP2C19 and disulfiram for CYP2E1.

mixtures of chemicals, it can be appreciated that prediction of the final outcome of metabolisms that are affected can be quite complex.

Listings such as in Table 2.4 do not indicate the mechanism of inhibition, which can be by three mechanisms (Figure 2.9). As a consequence of the body having only a limited number of P450s with which to oxidatively metabolize hundreds of drugs and xenobiotics, when inhibitory interactions do occur, the mechanism is often simple competition for limited amounts of enzyme (Figure 2.9a). In the

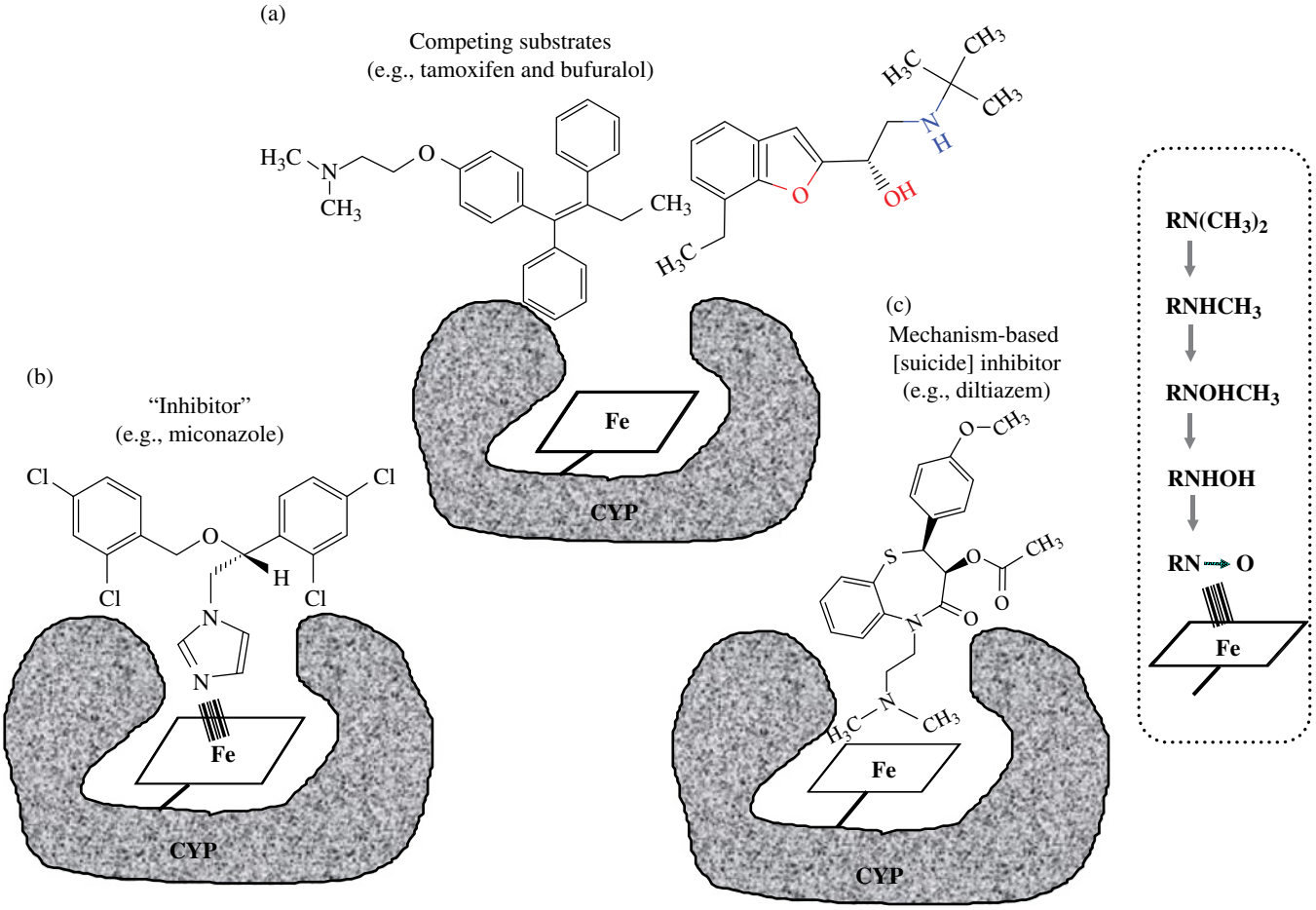


FIGURE 2.9 Competive, noncompetitive, and mechanism-based P450 inhibition.

unlikely event that the enzyme had similar affinity (K_m) and V_{max} values for two substrates, it would appear to the external observer that the metabolism of each was decreased ("inhibited") 50%. Note that the enzyme is not inhibited, and metabolism occurs at the same rate as when only one substrate was present. In most scenarios, the enzyme has greater affinity for one substrate over another leading to preferential metabolism of the one with the highest affinity, and apparent greater inhibition of the metabolism of the substrate (xenobiotic) with the lower affinity. The inhibition appears "competitive" since manipulating the substrate concentrations can almost negate the inhibition (e.g., by increasing the lower-affinity substrate concentration, it will outcompete the fixed [i.e., limited] concentration of the higher affinity substrate).

For xenobiotics that can interact strongly with part of the P450 in addition to the "substrate binding" site (e.g., the heme moiety), the affinity between xenobiotic and enzyme can be increased substantially, and can be so avid that entry into the active site and displacement of inhibitor by a substrate is virtually impossible. In this scenario, the inhibition of substrate metabolism appears to be noncompetitive (the inhibitor cannot be displaced by a substrate, however high its concentration is raised) (Figure 2.9b). Examples of such heme ligand inhibitors are antifungal agents (designed to inhibit fungal CYP-dependent ergosterol synthesis, a key component of fungal membranes) and include N-substituted imidazoles (e.g., clotrimazole, econazole, miconazole, butoconazole, ketoconazole, oxiconazole, sulconazole) and triazoles (e.g., fluconazole, voriconazole, itraconazole, terconazole).

There is great interest and concern about xenobiotics that demonstrate unusual inhibition characteristics: those where their P450 inhibition changes from competitive to noncompetitive upon metabolism. With metabolism, the xenobiotic gains the ability to interact at other than the substrate binding site, in some cases binding in a covalent manner to the protein or heme (or both), causing irreversible inhibition. These are mechanism-based or "suicide" inhibitors (Figure 2.9c). For mechanism-based inhibition, the cytochrome P450 inhibited is most likely to be the enzyme that created the reactive metabolite. Substrate promiscuity means that more than one P450 may be affected. This reaches an extreme with the metabolism of 1-aminobenzotriazole, which is catalyzed by most P450s and results in the formation of a heme adduct.

TABLE 2.5 Xenobiotics Forming Inhibitory Adducts With Cytochrome P450s

CYP	Inhibitory Adduct(s) (where delineated) and substrate
2A6	Protein adducts from benzylisothiocyanate and 8-methoxypsoralen
2B6	Protein adducts from bergamottin (grapefruit juice), ethynylestradiol, glabridin, phencyclidine, and tamoxifen and a heme adduct from N′,N″triethylenethiophosphoramide.
2C9	Protein adducts from tamoxifen and tienilic acid
2C19	Protein adduct from ticlopidine
2D6	Protein adducts from raloxifine, tamoxifen, and posssibly parathion
2E1	Adducts from benzylisothiocyanate, diallyl sulfide (garlic oil) and dichloroethylene
3A4	Adducts from bergamottin (grapefruit juice), ethynylestradiol (heme adduct), 8-geranyloxypsoralen, glabridin, parathion, raloxifine, and zafirlukast

TABLE 2.6 Cytochrome P450–Selective Reactions

CYP	Popular or "FDA Guidance" model reactions
1A2	Phenacetin O-deethylation, methoxyresorufin O-demethylation, ethoxyresorufin O-deethylation, **theophylline** N-demethylation, **caffeine** 3-N-demethylation, tacrine 1-hydroxylation.
2A6	Coumarin 7-hydroxylation, nicotine C-oxidation
2B6	**Efavirenz** hydroxylase, bupropion hydroxylation, propofol hydroxylation, *S*-mephenytoin N-demethylation, 7-ethoxy-4-trifluoromethylcoumarin O-deethylation.
2C8	Paclitaxel 6-hydroxylation, amodiaquine N-deethylation, **rosiglitazone** p-hydroxylation
2C9	Diclofenac 4-hydroxylation, **tolbutamide** methyl-hydroxylation, ***S*-warfarin** 7-hydroxylation, flurbiprofen 4′-hydroxylation, phenytoin 4-hydroxylation, luciferin-H hydroxylation.
2C19	*S*-mephenytoin 4-hydroxylation, **omeprazole** 5-hydroxylation, fluoxetine O-dealkylation.
2D6	Bufuralol 1′-hydroxylation, **dextromethorphan** O-demethylation, debrisoquine 4-hydroxylation, 7-methoxy-4-(aminomethyl)coumarin-O-demethylation.
2E1	**Chlorzoxazone** 6-hydroxylation, 4-nitrophenol 3-hydroxylation, lauric acid 11-hydroxylation, aniline 4-hydroxylation.
3A4/5	**Midazolam** 1-hydroxylation, testosterone 6β-hydroxylation, erythromycin N-demethylation, dextromethorphan N-demethylation, **triazolam** 4-hydroxylation, terfenadine C-hydroxylation, nifedipine oxidation, 7-benzyloxyquinoline debenzylation.

Drugs shown in **bold** are substrates suggested (FDA Guidance for Industry Publication, September 2006) for use for *in vivo* inhibition studies. Other possible drugs that can be used include repaglinide for CYP2C8, esoprazole, lansoprazole, and pantoprazole for CYP2C19, desipramine and atomoxetine for CYP2D6, and buspirone, felodipine, lovastatin, eletriptan, sildenafil, and simvastatin for CYP3A4/5.

Other CYP mechanism–based inhibitions are provided in Table 2.5. A considerable number of drugs and other xenobiotics are able to form "nitroso" quasi-irreversible metabolic intermediate complexes with P450s. Those demonstrated with CYP3A4 include amprenavir, clarithromycin, desipramine, diltiazem (shown in Figure 2.9c), erythromycin, fluoxetine, fluvoxamine, indinavir, nelfinavir, nicardipine, nortriptyline, ritonavir, verapamil, tamoxifen, and troleandomycin. Also, quasi-irreversible metabolic intermediate complexes are formed with methylenedioxyphenyl (benzodioxole)-containing compounds. The methylenedioxyphenyl moiety is present in a few drugs (e.g., paroxetine), and in many plant and botanical preparations (e.g., berberine and hydrastine in goldenseal, methysticin in kava, myristicin in parsley and nutmeg, piperine in black pepper, and, previously available, safrole in sassafras). Piperonyl butoxide, a synthetic methylenedioxyphenyl (benzodioxole) compound, is used as a pesticide synergist, especially for the pyrethroids. Propyl isome, sesamex and sesamolin, and piprotal are additional benzodioxoles and enter the environment for similar reasons.

The ability of and mechanism by which drugs and xenobiotics inhibit P450s is most often established with *in vitro* assays using isolated hepatic microsomes or reconstituted heterologously expressed enzymes and model substrates and reactions deemed to be selective for a single enzyme (Table 2.6). In some cases, chemical reporter substrates, usually yielding fluorescent metabolites, have been created specially for this purpose.

The consequences of P450 DDI inhibition will depend on the activating/inactivating nature of the reaction catalyzed. If CYP-dependent metabolism of a xenobiotic results in inactivation, inhibitors will likely increase the pharmacological or toxicological effect. In those instances where metabolism results in bioactivation of a drug or xenobiotic, inhibitors will decrease the pharmacological or toxic effect. Of note in this respect, numerous anticancer agents are prodrugs and require P450-dependent "activation" for activity. They include cyclophosphamide (2B6 and 2C19), dacarbazine (1A1, 1A2, and 2E1), ifosfamide (2B6), procarbazine (2B6 and 1A2), and thiotepa (3A and 2B6).

Dehydrogenases The major enzymes utilizing a dehydrogenation mechanism for the oxidation of xenobiotics are generally considered as alcohol (ADHs) and aldehyde dehydrogenases (ALDHs). Dehydrogenase-catalyzed oxidations do not directly involve molecular oxygen. The oxidation of the xenobiotics by dehydrogenases occurs through electron transfer to a pyridine nucleotide, usually NAD^+. While the ADHs are cytoplasmic, ALDHs are found in mitochondria, the endoplasmic reticulum, and cytoplasm. The two enzymes work in concert to convert alcohols to aldehydes and onward to carboxylic acids. For ethanol, this results in its complete oxidation to acetate, passing through the potentially more toxic acetaldehyde.

Alcohol Dehydrogenase ADHs are divided into five classes, of which I, II, and III are of greatest interest in the hepatic metabolism of xenobiotics. Class I enzymes preferentially oxidize small aliphatic alcohols, such as ethanol, while the Class II and the Class III enzymes oxidize larger and longer-chain aliphatic and aromatic alcohols. The classes separate similarly with respect to effects of pyrazole and 4-methylpyrazole, which only inhibit Class I enzymes. The Class I enzymes have a K_m value well below the blood concentration achieved in ingesting moderate amounts of alcoholic beverages and so ethanol metabolism is generally considered to show zero-order kinetics (fixed metabolism per unit time, independent of blood levels). ADHs are active as dimers. The Class II enzyme (ADH4) comprises two pi subunits, and the Class III enzyme (ADH5) comprises two chi subunits. The Class I enzymes can be hetero- as well as homodimers and while ADH1C is a homodimer of gamma subunits, ADH1B can comprise two beta or one beta and one gamma subunits, and ADH1A can be homo alpha, alpha and beta, or alpha and gamma subunits. Subunit composition of Class I enzymes can be important since those containing beta subunits, especially those with one (beta2) of the three allelic variants of the beta subunit (which each differ by a single amino acid), are more active in metabolizing ethanol, and are responsible for the rapid formation of acetaldehyde in East Asian populations. Not surprisingly, higher V_{max} ADHs are underrepresented in the alcoholic population, and it is believed the basis of this is that faster production of toxic acetaldehyde and the adverse effects it produces dissuades heavy ethanol ingestion.

In the realm of clinical toxicology, substrate overlap in ADHs is utilized in the treatment of methanol and ethylene glycol poisoning. The oxidation of these compounds via their aldehydes to formic and oxalic acid can be reduced by administration of a competing substrate, ethanol. The ADH-dependent formation of acrolein ($CH_2=CHCHO$) from allyl alcohol ($CH_2=CHCH_2OH$) forms the basis of its hepatotoxicity. In laboratory animal studies, this toxicity has been found to be localized to the periportal region of the hepatic lobule.

Aldehyde Dehydrogenase Almost 20 ALDHs are present in humans and a similar number appear present in other mammalian species, suggesting a survival mechanism against the accumulation of the reactive aldehyde group in chemicals of widely different structure. While some are located in cytosol and the endoplasmic reticulum, ALDH2, the major form responsible for the metabolism of simple aldehydes, such as acetaldehyde, is localized to the mitochondria. Relevant to the observations on ADH variants (discussed earlier), East Asian populations have a higher prevalence of a variant form of ALDH2, one in which a G to A base change results in a Glu487Lys conversion, resulting in an enzyme with little or no activity. With a higher quantity of enzymes that rapidly produce acetaldehyde and poorly remove it, such populations rapidly display the easily recognizable flushing syndrome (catechol-induced vasodilation) upon ingestion of only limited quantities of ethanol. Together with flushing, other effects of acetaldehyde, notably headache, nausea, and vomiting ("hangover"), can be elicited by ethanol ingestion when ALDH activity has been inhibited with disulfiram (Antabuse). This forms the basis of part of the aversion therapy used to help alcoholics change their lifestyle.

Monoamine Oxidases MAOs catalyze an oxidative deamination reaction. The reaction proceeds by electron transfer to a flavin group (FAD), which upon the interaction of $FADH_2$ with oxygen produces H_2O_2, which in turn is converted by the enzyme catalase to water and oxygen. The initial hydrogen abstraction come from the α-carbon, which explains why compounds (drugs) that contain a methyl group attached to that carbon (e.g., amphetamines) are not metabolized by this enzyme. The ability to also abstract hydrogen from the terminal nitrogen of certain hydrazines, notably phenelzine, leading to the formation of a reactive diazine that covalently adducts to and inhibits the enzyme, is the mechanism of action of this antidepressant drug. Inhibition of MAOs prevents metabolism and allows the accumulation of neurotransmitters necessary for maintaining mental balance. There are two MAOs (A and B) each coded for by similar but distinct genes on the X chromosome. The enzymes, both expressed in mitochondria, display some differences in substrate and inhibitor selectivity. For instance, MAO-A preferentially metabolizes 5-hydroxytryptamine (serotonin) and norepinephrine and is inhibited through a mechanism-based (suicide) pathway by chlorgyline. MAO-B preferentially metabolizes β-phenethylamine, benzylamine, and 1-methyl-4-phenyl-1,2,5,6-tetrahydropyridine (MPTP) and is suicide-inhibited by selegiline (L-deprenyl).

In addition to naturally occurring catecholamines (dopamine, epinephrine, norepinephrine) and tryptophan derivatives (serotonin), there are many drugs that are substrates for MAOs. These include triptans (sumatriptan, zolmitriptan, rizatriptan), doxylamine, haloperidol, and primaquine.

Naturally occurring substances such as tyramine (present in many fermented, marinated, pickled, aged, smoked foods, and certain fruits from the decarboxylation of tyrosine) are also substrates and hence the warning about ingestion of such foodstuffs when MAO activity is inhibited with antidepressant therapy since their accumulation can result in the release of stored monoamine and result in a hypertensive crisis.

Phase I Hydrolysis Hydrolysis reactions are catalyzed by carboxyl and cholinesterases with the former localized in tissues rather than plasma. In addition to the carboxyl and cholinesterases, which contain a serine residue in their catalytic site (the site of inhibition by organophosphorus compounds), another group of enzymes (paraoxonases) that catalyze hydrolysis reactions, particularly of the organophosphates, utilize a sulfhydryl residue in their active site. Carboxyl and cholinesterases are inhibited by organophosphates because the phosphorous–oxygen bond formed at the active site serine residue is resistant to cleavage by water. However, it can be dephosphorylated by pralidoxime (2-PAM), thereby regenerating enzyme activity.

The hydrolysis reactions can be summarized as for esters, $RCOOR \rightarrow RCOOH + ROH$, and for amides, $RCONHR \rightarrow RCOOH + RNH_2$. It should be noted that the hydrolysis reaction produces two entities each with a chemically reactive functional group, and both are therefore suitable for conjugation, if the metabolites are not first eliminated as Phase I products. For similar chemicals, the hydrolysis of amides (e.g., procainamide) is slower than of esters (e.g., procaine). Drugs hydrolyzed come from many pharmacological classes. Examples include local anesthetics (e.g., cocaine, lidocaine, procaine), narcotics and analgesics (e.g., aspirin, heroin, indomethacin, meperidine), parasympatholytics, muscle relaxants and vasodilators (e.g., pancuronium, succinylcholine), neuromuscular blocking agents (e.g., atracurium), antiarrhythmics (e.g., lidocaine, procainamide), antibiotics (e.g., chloramphenicol), CNS stimulants (e.g., methylphenidate), and drugs used to treat dyslipidemia (e.g., clofibrate). While hydrolysis often results in loss of pharmacological activity, this is not always the case. Some prodrugs are activated by the reaction. Examples include lovastatin (hydrolyzed to lovastatin β-hydroxyacid, the active HMG CoA reductase inhibitor), irinotecan (CPT-11) (hydrolyzed to SN38, the topoisomerase I inhibitor), and enalapril (hydrolyzed to the active angiotensin-converting enzyme (ACE) inhibitor).

There are two major carboxylesterases (CES1 and CES2) with CES1 showing a greater activity for molecules with a smaller alcohol "end" and CES2 showing a preference for a molecule with a smaller acid "end." CSE1 is responsible for the bioactivation of many prodrug ACE inhibitors (e.g., benzapril, enalapril, and quinapril) also in the bioactivation of oseltamivir (Tamiflu). Meperidine is inactivated by CSE1. CES2 substrates include 6-acetylmophine, aspirin (acetylsalicylic acid), cocaine benzoyl ester, heroin, irinotecan, and oxybutynin.

Pseudo- or butyryl-cholinesterase is a tetrameric enzyme found in the plasma and was one of the earliest drug response variations attributed to metabolism. Some patients showed exaggerated muscle relaxation response to the neuromuscular blocking agent succinylcholine, and the incidence varied with populations. It was found that the enzyme in the abnormal responders had reduced activity toward the drug because of a lowered affinity (higher K_m). Upon investigation, the variant enzyme was also more resistant to the inhibitory effects of dibucaine on benzoylcholine hydrolysis. Dibucaine is a local anesthetic that is not a substrate. This phenomenon gave rise to a test for the variants (dibucaine number = % inhibition under specified benzoylcholine and dibucaine concentrations; normal ~78, variant ~4). The altered activity was eventually traced to an aspartate to glycine substitution at amino acid residue 70.

The frequency of the abnormal psuedocholinesterase gene varies with populations. It is ~2% in British, Greek, Portuguese, North African, Jewish, and some Asiatic groups, rare in Africans, Australian aborigines, Filipinos, and Asians other than the Japanese, and essentially absent in Eskimos, South American Indians, and the Japanese. Substrates in addition to succinylcholine include cocaine, methylprednisone acetate, heroin, isosorbide diaspirinate, procaine, and tetracaine.

Phase I Reductions Reduction reactions can be catalyzed by many different groups of enzymes, including some that are normally associated with oxidation reactions (e.g., cytochrome P450). Reductive dehalogenation of carbon tetrachloride and halothane (chlorine and bromine loss, respectively) by cytochrome P450 is such an example, and this one-electron reduction by cytochrome P450 produces a carbon-centered radical that sets the stage for lipid peroxidation and the membrane damage associated with these potentially hepatotoxic molecules. In many cases, reduction reactions occur most readily in low oxygen tension environments. For instance, reduction of nitro (e.g., chloramphenicol) and azo (e.g., prontosil) groups (to amines) is often catalyzed by bacterial enzymes in the oxygen-depleted environment present in the lower regions of the intestinal tract. The reduction of prontosil, $RN=NR' \rightarrow RNH_2 + R'NH_2$, is of historical significance in that one of the amine products (sulfanilamide) rather than the azo dye itself proved to be the active antibacterial agent, and therefore prontosil was active in animals, but not *in vitro*. Thus was born the sulfonamide antibacterials still in wide use today.

For some compounds, reduction of cytochrome P450 and FMO-generated oxidation products can occur. This is commonly seen with sulfoxide and N-oxide metabolites. Reduction of aldehydes generated by dehydrogenases also

can occur, a reaction catalyzed by aldo-keto reductases present in the cytosol, many of which are themselves categorized as dehydrogenases.

Of special note among reduction reactions is the two-electron reduction of quinones to hydroquinones by an enzyme originally known as DT-diaphorase, now NAD(P)H-quinone oxidoreductase, and abbreviated as NQO. The two-electron reduction is seen as being less harmful than a one-electron reduction to a reactive semiquinone species, which, in addition to being intrinsically reactive, can generate toxic reactive oxygen species (ROSs) from molecular oxygen. Expression of this enzyme together with several other "protective" enzymes is acutely sensitive to the redox status of the cell, through the Nrf2–Keap 1 pathway and the possession of an antioxidant response regulatory element (ARE) (discussed later in the Xenobiotic-Metabolizing Enzyme Induction section).

Phase I½ Epoxide Hydrolase The metabolism of epoxides by the addition of water as catalyzed by epoxide hydrolases is a reaction that is difficult to classify. In most instances it follows Phase I oxidation by cytochrome P450 of xenobiotics (aromatic hydrocarbons and aliphatic alkenes) to an epoxide; it does confer increased hydrophilicity and, most importantly for toxicity considerations, decreased chemical reactivity, but the enzyme is not termed a transferase as are most Phase II enzymes, and water is hardly an activated cosubstrate. The reaction is catalyzed by epoxide hydrolases; for xenobiotic epoxides there are two, one located in the endoplasmic reticulum (microsomes; mEH) and another found in the cytosol (soluble; sEH). The enzymes, presumably related to their epoxide detoxication role, are found in those tissues and cells that express higher levels of cytochrome P450. Although the two epoxide hydrolases have some differing substrate preferences, both catalyze the trans addition of water to arene oxides and alkene epoxides to form vicinal trans diols. Important in the reactive center of the enzymes and key to the reaction mechanisms are two acidic amino acid residues, a histidine and a tyrosine, even though the two enzymes show little overall sequence homology.

Phase II Conjugations All conjugations involve the participation of enzymes characterized as transferases and, without exception, they occur in multiple forms, often with differing substrate selectivities. All conjugation reactions, except with glutathione, involve the participation of energy-rich or activated cosubstrates. Thus for glucuronidation catalyzed by glucuronosyltransferases, the enzyme "transfers" the glucuronic acid moiety from uridine 5′-diphospho-glucuronic acid (UDPGA) to the xenobiotic or Phase I metabolite of the xenobiotic. For sulfate conjugation, the cosubstrate is 3′-phosphoadenosine-5′-phosphosulfate (PAPS), for acetylation the cosubstrate is acetyl–coenzyme A, and for methylation the cosubstrate is *S*-adenosylmethionine (SAM). Amino acid conjugation is a little different in that it is the xenobiotic or the Phase I metabolite of the xenobiotic that is the "activated" cosubstrate. In terms of the involvement of the conjugation reactions in the metabolism of all xenobiotics, glucuronidation, sulfation, and glutathione conjugations are among the most frequently encountered.

Glucuronidation UDP–glucuronosyltransferases (abbreviated UGTs) are enzymes located in the endoplasmic reticulum (microsomes *in vitro*) that catalyze the transfer of glucuronic acid from UDPGA to, most often, a hydroxyl (phenol) group (e.g., acetaminophen, bisphenol A) to form an ether glucuronide, a carboxyl group (e.g., valproate, naproxen) to form an alkali labile ester glucuronide, or an amine or substituted amine group (e.g., sulfathiazole, lamotrigine) to form an *N*-glucuronide. The UDPGA is generated from glucose-1-phosphate, which is readily derived from glycogen, the abundant carbohydrate supply present in the liver.

Xenobiotic glucuronide formation does not involve the acidic functional group of glucuronic acid so the conjugate retains the carboxylic acid and therefore the extensive ionized character at physiological pHs (Figure 2.10, also Figure 2.5). This provides the dramatic enhancement of water solubility and passive excretability. Glucuronides (anions) are also actively secreted into bile and into urine by transporters. The range of compounds excreted as glucuronides is exceedingly diverse and encompasses endogenous compounds, therapeutic drugs, and numerous compounds present in botanical preparations (Table 2.7). For many plant products, the compound glucuronidated must first be generated. Phytoestrogens, for example, are polycyclic phenols that occur in certain plants as beta-glycosides and are hydrolyzed to the aglycone in the intestine before being absorbed and subsequently glucuronidated in the liver. A similar scenario occurs with the plasticizer phthalate esters encountered in the diet, although for these compounds, intestinal hydrolysis is followed by oxidation prior to glucuronidation.

UGTs responsible for xenobiotic conjugations occur in multiple forms, and in two families that are quite distinct in gene structure and chromosomal location, but may exhibit overlapping substrate (aglycone) selectivity. For example, the O-glucuronidation of hydroxyl groups of the tamoxifen metabolites, 4-hydroxytamoxifen and endoxifene (4-hydroxy *N*-desmethyl tamoxifen), is catalyzed by UGT1A10 and UGT2B7. For either family, the 5′ exon(s) (toward the N-terminal end of the protein) appears to code for aglycone recognition while the 3′ (C-terminal end of the protein) codes for recognition of UDPGA. The gene structure of the UGT1A family of enzymes (UGT1A1 and UGT1A3 through UGT1A10), located on chromosome 2, is such that independent first exons, each with their own transcription start site, are spliced to a common set of exons 2–5. Noting

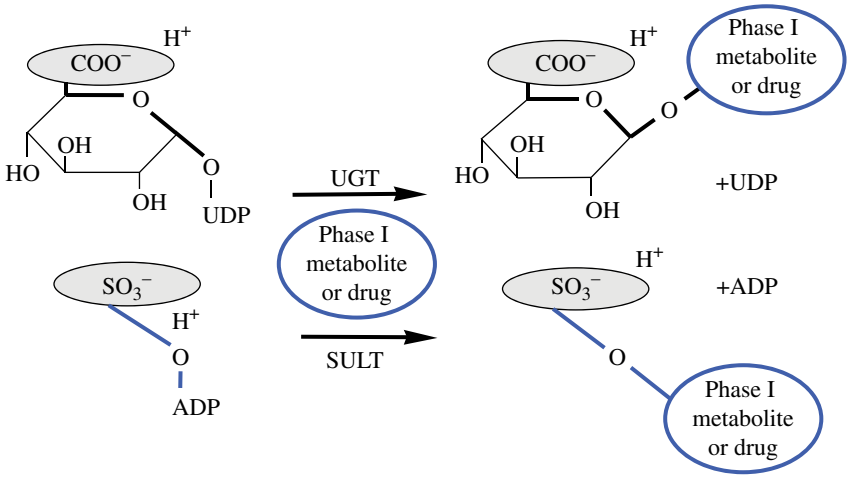


FIGURE 2.10 Commonality of glucuronidation and sulfation.

TABLE 2.7 Xenobiotics Metabolized to Glucuronides

UGT	Xenobiotic[a]
1A1	Acetaminophen, atorvastatin, **bilirubin**, buprenorphine, carvedilol, estradiol, estriol, ethinylestradiol, etoposide, ezetimibe, febuxostat, fluvastatin, gemfibrozil, irinotecan, levothyroxine, losartan, lovastatin, ketoconazole, morphine, naltrexone, naphthol, raloxifene, raltegravir, rosuvastatin, simvastatin, SN-38, telmisartan, thyroxine, triiodothyronine
1A3	Amitriptyline, atorvastatin, buprenorphine, chlorpromazine, clofibrate, clozapine, cyproheptadine, diclofenac, diphenhydramine, doxepin, ezetimibe, febuxostat, fenoprofen, fluvastatin, gemfibrozil, 4-hydroxytamoxifen, ibuprofen, imipramine, ketoprofen, losartan, lovastatin, loxapine, nalorphine, naloxone, naltrexone, naproxen, promethazine, rosuvastatin, simvastatin, valproate
1A4	Amitriptyline, androsterone, asenapine, chlorpheniramine, chlorpromazine, clozapine, cyclobenzaprine, cyproheptadine, dapsone, desmethylclozapine, diphenhydramine, doxepin, 4-hydroxytamoxifen, **imipramine**, irinotecan, ketotifen, lamotrigine, loxapine, meperidine, nicotine, olanzapine, progestins, promethazine, pyrilamine, retigabine, trifluoroperazine
1A6	Acetaminophen, entacapone, ketoprofen, naproxen, nefazodone, **serotonin**, valproate
1A8	Morphine, mycophenolate, **propofol**, raloxifene
1A9	Acetaminophen, bumetanide, clofibrate, dapsone, diclofenac, diethylstilbestrol, diflusinal, entacapone, estradiol, estrone, febuxostat, fenofibrate, furosemide, gemfibrozil, 4-hydroxytamoxifen, ibuprofen, ketoprofen, labetalol, mycophenolate, naproxen, nicotine, propofol, propranolol, SN-38, retinoic acid, thyroxine, tolcapone, triiodothyronine, zidovudine
1A10	Estradiol, estrone, furosemide, raloxifene, simvastatin, SN-38
2B4	Bile acids, carvedilol, codeine, estriol
2B7	Almokalant, atorvastatin, buprenorphine, carvediolol, chloramphenicol, clofibric acid, codeine, cyclosporine, diclofenac, entacapone, epirubicin, estriol, febuxostat, fenofibrate, fluvastatin, gemfibrozil, hydromorphone, ibuprofen, ketoprofen, lamotrigine, lorazepam, methadone, **morphine** (both phenol 3-OH and alcohol 6-OH), mycophenolate, nalbuphine, nalmefene, nalorphine, naloxone, naltrexone, naproxen, nicotine, oxymophone, oxazepam, oxycodone, all-trans-retinoic acid, simvastatin, tacrolimus, temazepam, valproic acid, zidovudine
2B15	Bisphenol A, entacapone, estrogens, ezetimibe, gemfibrozil, lorazepam-S, oxazepam-S, testosterone, tolcapone, 17β-hydroxyandrogens

[a]Xenobiotics are listed alphabetical per enzyme. Note that there is considerable inter-enzyme overlap. Compounds shown in **bold** are suggested probe substrates for the enzyme.

the independent transcription start sites is important in understanding a UGT1A1 polymorphism (UGT1A1*28) where the normal promoter region (TATA box) has additional base insertions, that is [A(TA)$_7$TAA instead of A(TA)$_6$TAA]. The extra bases lead to low expression, but only of UGT1A1, not UGT1A3-UGT1A10. UGT1A1 is important in the metabolism of bilirubin, and low expression can lead to hyperbilirubinemia (Gilbert's disease), but UGT1A1 is also involved in the glucuronidation of xenobiotics. A reduced ability to conjugate the phenolic metabolite (SN-38) generated by ester hydrolysis (by CES2) of the anticancer drug irinotecan leads to increased GI toxicity of this chemotherapeutic agent. As with some of the cytochrome P450s, the frequency of this polymorphism varies with

TABLE 2.8 Drugs and Endogenous Compounds Metabolized To Sulfate Conjugates

Xenobiotic	Sulfotransferase(s)[a]								
Hydroxyarylamines	1A1	1A2							
Acetaminophen	1A1		1A3						
Apomorphine	1A1		1A3						
Minoxidil	1A1								
Genistein, naringenin	*1A1*						1E1		
4-Hydroxytamoxifen	1A1						*1E1*		
'Iodothyronines'	1A1			1B1					
Thyroxine (T_4)				1E1					
Catechols	1A1	1A2	1A3	1B1					
Dopamine	1A1		1A3	1B1					
Epinephrine	1A1		1A3						
Norepinephrine, 5-hydroxytryptamine			1A3						
Tyramine, albuterol, isoproterenol, dobutamine			1A3						
N-Hydroxyacetylaminofluorene		1A2			1C2	1C4			
17β estradiol (E2), estrone	1A1						1E1	2A1	
17α ethinylestradiol	1A1						1E1	2A1	
Diethylstilbestrol (DES)	1A1						*1E1*		
Cholesterol								2A1	2B1
Pregnenolone, dehydroepiandrosterone (DHEA)							1E1	2A1	2B1
17α hydroxypregnenolone									2B1
Testosterone, androsterone cortisol, 'bile acids'								2A1	
Dihydrotestosterone									2B1
Etiocholanolone								2A1	2B1

[a]SULT enzymes shown in italics are minor contributors to conjugation.

ethnicity; in this case, it is threefold higher in Caucasians than in Japanese and Chinese populations. Because the polymorphism under discussion is UGT1AI-specific, it should be noted that the N-glucuronidation of tamoxifen and 4-hydroxytamoxifen, catalyzed by another UGT1A family member, UGT1A4, is unaffected.

In contrast to the UGT1A family, each member of the UGT2B family (UGT2B4, UGT2B7, UGT2B10, UGT2B15, and UGT2B17) and located on chromosome 4 has an independent gene, consisting of six exons. As with the UGT1A family, polymorphisms also exist among UGT2B family members, but the consequences for xenobiotic substrates have not been extensively studied. Two of the earliest identified polymorphisms showed greater frequency in Asian than in Caucasian populations.

Sulfation Sulfotransferases (abbreviated SULTs), in contrast to UGTs, are located in the cytoplasm, but similar to UGTs are involved in the conjugation of hydroxyl groups (Figure 2.3). Sulfate conjugation is also a metabolic route for some arylamines. The PAPS cosubstrate from which the sulfate group is transferred is generated from ATP and inorganic sulfate, the latter mostly derived from the sulfur-containing amino acids cysteine and methionine. PAPS is somewhat limited in availability, and compared to glucuronidation, sulfation may be considered a lower capacity pathway. As such, at higher concentrations of xenobiotics or Phase I metabolites containing a phenolic functional group, the balance of drug metabolism may shift toward a greater percentage being subject to glucuronidation. Sulfate conjugates are extensively ionized at physiological pH (Figure 2.10) and therefore, like glucuronides, are easily eliminated, either by glomerular filtration or by active transport in the proximal tubule of the kidney.

As with UGTs, SULTs are present as multiple enzymes and also show considerable overlap in substrate selectivity. In addition to recognizing xenobiotics, many SULTs are involved in the metabolism of endogenous compounds, especially steroid hormones and neurotransmitters, and are often referred to by their endogenous substrate preference. Thus 1A3 is commonly identified as the catecholamine/dopamine form, 1B1 as the thyroid hormone form, 1E1 as the estrogen form, 2A1 as the DHEA (dehydroepiandrosterone) form, 2B1a/2B1b as the 3β hydroxysteroid form, even though, as can be seen in Table 2.8, there is a high degree of substrate overlap.

Glutathione Conjugation Glutathione (*S*-)transferases (abbreviated GSTs) constitute one if not the major component of hepatic cytosolic protein. They are important for sequestering reactive (toxic) electrophilic metabolites generated in the cell (e.g., epoxides generated during cytochrome P450–dependent oxidation). Glutathione (γ-glutamyl–cysteinyl–glycine, GSH) is a major cellular nucleophile (at

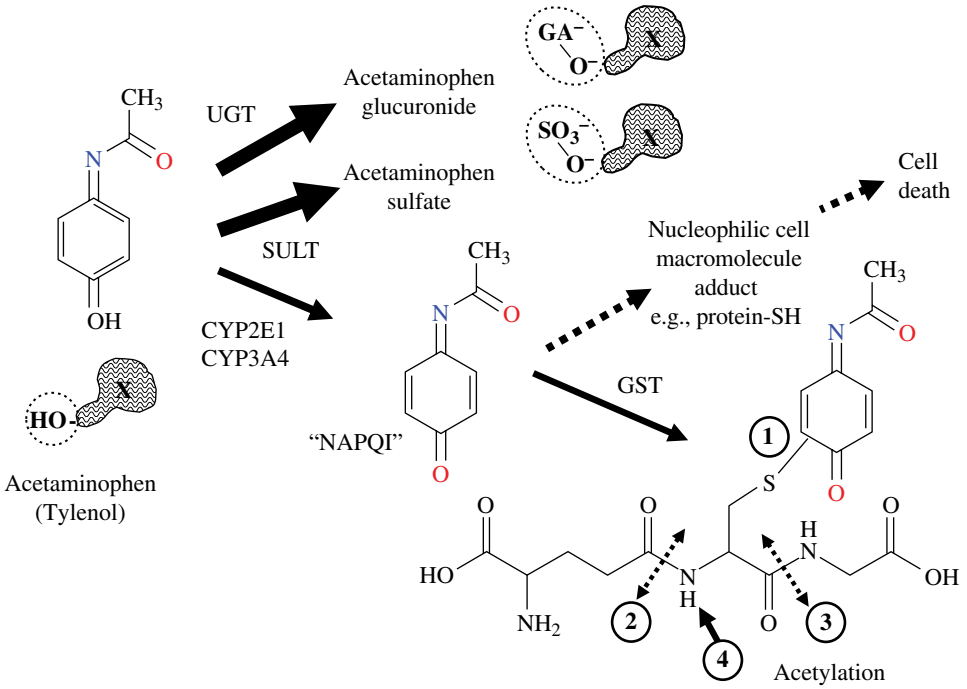


FIGURE 2.11 Acetaminophen metabolism and toxicity.

the –SH group of the cysteine residue) and is sufficiently reactive that for some compounds, glutathione conjugation can occur without the participation of enzymes. The glutathione cosubstrate is abundantly present in the liver (>5 mM).

Like the other common Phase II xenobiotic-metabolizing enzymes, GSTs are present as multiple forms. The soluble GSTs fall into seven classes (alpha, mu, pi, theta, zeta, omega, sigma) and in humans there are four alpha forms, five mu forms, and two theta forms. Isozyme expression shows considerable interorgan variation. GSTs function as dimeric proteins, and, in common with all xenobiotic-metabolizing enzymes, have overlapping substrate selectivity. In contrast with glucuronic acid and sulfate conjugates, glutathione conjugates rarely appear as such during excretion, but emerge as secondary metabolic products termed mercapturic acids (acetylated cysteine derivatives). They are formed by sequential removal of glutamate and then glycine from the glutathione portion followed by acetylation (catalyzed by *N*-acetyltransferase; see later) of the amino group of the residual cysteine. The critical need for glutathione conjugation can be demonstrated for the common analgesic, acetaminophen (Figure 2.11). In the United Kingdom, acetaminophen hepatotoxicity is the leading cause of acute liver failure. Acetaminophen at low doses is extensively conjugated by sulfation and glucuronidation. Small amounts, more at higher doses, escape these conjugations and undergo CYP2E1 (and a minor contribution from CYP3A4)-dependent oxidation to *N*-acetyl-p-benzoquinone imine (NAPQI). This reactive metabolite can interact with nucleophilic sulfhydryl groups, and, if with the –SH group of glutathione, forms (with "amino acid trimming" modification) an excretable metabolite. In the event of insufficient glutathione or GST activity, the interaction with protein –SH groups can disrupt protein function, and if a critical function, can result in cell death.

Other Conjugations There are two conjugations, acetylation and methylation, that do not contribute to enhanced excretability through an increase in water solubility. This was well demonstrated with acetylation of some early sulfonamides where the acetylated metabolites were sufficiently less water soluble than the parent drug that they precipitated as crystals in the urine and renal cells and resulted in kidney tissue damage. Rather than facilitating excretion, acetylation and methylation are best viewed as masking reactive centers.

Acetylation Acetylations are catalyzed by *N*-acetyltransferases. There are two such enzymes, NAT1 and NAT2, and the latter is the form responsible for the bulk of the acetylations occurring in the liver. Acetyltransferases are located in the cytosol. The acetyl–coenzyme A (a thioester) cosubstrate for the conjugation is readily available from intermediary metabolism. The reaction proceeds through an acetylated enzyme intermediate (releasing coenzyme A).

A major "claim to fame" of NAT2 is the presence of a genetic polymorphism, which was among the earliest identified for the metabolism of xenobiotics, and from where emerged the field of pharmacogenetics. The polymorphism gave rise to the identification of "fast" and "slow" acetylator phenotypes, particularly as it related to isoniazid, a drug

TABLE 2.9 Methyltransferase Nomenclature and Characteristics

Acceptor Atom	Enzyme Abbreviation	Subcellular Localization	Example Substrates
O	Phenol (POMT)	Microsomal	Phenols (not catechols)
	Catechol (COMT)	Cytosolic	Epinephrine, norepinephrine, dopamine, L-dopa, α-methyldopa, catechol (A ring), estrogens, Phase I catechol metabolites
N	Histamine (HNMT)	Cytosolic	Imidazole ring of histamine and related compounds
	Nicotinamide (NNMT)		Pyridines/indoles, nicotine, tryptophan, serotonin
	Phenylethanolamine (PNMT)		Norepinephrine (to epinephrine)
S	Thiopurine (TPMT) cytosolic		6-Mercaptopurine, 6-thioguanine, azathioprine
	Thiol (TMT)	Microsomal	Aliphatic–SH: disulfiram, D-penicillamine, captopril

used worldwide in the chemotherapy of tuberculosis. The noticed variations in the frequency of the phenotypes in different ethnic populations (compared to Caucasian populations, the percentage of slow metabolizers is higher in Middle Eastern populations and lower in Asian populations) was also influential in the consideration of population differences in drug therapeutic effectiveness and drug, and eventually environmental and occupational chemical toxicity (toxicogenetics). Numerous NAT2 allelic variants have been identified. The most common "slow metabolizer" mutation in Caucasians (the basis of the NAT2*5 allele) is T341C (thymine to cytosine) giving Ile114Thr (isoleucine to threonine). This change decreases the enzyme V_{max}, with no change in K_m (affinity). In the Chinese, the most common mutation is the NAT2*7 allele where G857A (guanine to adenine) results in a Gly286Glu (glycine to glutamate) substitution, which decreases enzyme stability.

Drugs acetylated include *p*-aminobenzoic acid, aminoglutethimide, *p*-aminosalicylic acid, dapsone, hydralazine, isoniazid, procainamide, and sulfamethoxazole. In terms of isoniazid toxicity in relation to acetylation, early studies found that with prolonged dosing, slow acetylators showed a higher incidence of peripheral neuritis, fast acetylators, a higher incidence of liver toxicity. With higher plasma levels of isoniazid arising from its longer half-life in slow acetylators, the neuropathy was explained as arising from a chemical reaction (hydrazone formation) between it and pyridoxine (vitamin B6). The higher incidence of hepatic toxicity in fast acetylators was thought to be due to the more rapid formation of acetylisoniazid that underwent hydrolysis to acetylhydrazine, which could be metabolized (by CYP2E1) to acetyldiazine and thence an acetylcarbonium ion that could form covalent adducts with cellular components and result in cell death. Subsequent studies have associated liver toxicity with slow acetylator status, in which scenario it is the slower conversion of acetylhydrazine by a second acetylation reaction by NAT2 to diacetylhydrazine that allows greater diversion to the P450 oxidation intoxication pathway.

Methylation Methylations are catalyzed by *O*-, *N*-, and *S*(thio)-methyltransferases and each is present in multiple forms. Methyltransferases are often most conveniently referred to by their substrate preference (which in many cases are endogenous compounds) followed by their acceptor atom, for example, catechol O (oxygen) methyl transferase (Table 2.9). The SAM cosubstrate is readily derived from ATP and methionine.

Thiopurine methyltransferase (TPMT) has been the subject of much clinical investigation because of the toxicity of the drugs that it is able metabolize and the presence of activity-reducing polymorphisms. TPMT activity shows a trimodal distribution with <1% of the population showing low activity (low/low), approximately 11% intermediate activity (high/low), and approximately 88% high activity (high/high). 6-Mercaptopurine (Purinethol) is an antineoplastic agent used principally in the maintenance therapy of acute lymphocytic leukemia, and lower than normal metabolism causes a higher incidence of bone marrow depression (myelotoxicity) from "standard" dosing. Low activity is more common in Caucasian than in Asian populations. Catechol *O*-methyltransferase is unusual in that it exists in two forms, but both arise from a single gene (i.e., from two transcription start sites). The larger form with approximately 50 additional amino acids contains a membrane binding segment and is predominantly expressed in the brain whereas the smaller protein exists in the cytosol and is extensively expressed in the liver and kidney. Low activity polymorphism(s) exist but studies investigating whether these are linked to levodopa pharmacokinetic differences have shown both positive and negative support. Some studies have linked an increased risk of breast cancer to low activity, possibly linked to an enzyme activity toward catechol estrogens generated from estrone and 17β-estradiol.

Amino Acid Conjugation Amino acid conjugation is a reaction that provides little in the way of increased water solubility, and is probably also best regarded as a "masking" reaction. The conjugation takes place in two steps, an ATP-dependent formation by acyl–coenzyme A synthase of an acylcoenyzme A thioester on the carboxyl group of the xenobiotic (e.g., benzoic acid), followed by coenzyme A displacement by an amino acid, most often glycine or

glutamine, catalyzed by an *N*-acyltransferase. Glycine conjugates are known as hippurates.

Xenobiotic-Metabolizing Enzyme Induction

Enhanced activity can result from increased catalytic efficiency of existing enzyme (activation), but for xenobiotic-metabolizing enzymes this is relatively rare. Enhanced enzyme activity is most often the result of an increased synthesis of enzyme, a process termed induction. Because induction arises from the synthesis of a new enzyme, the increase in enzyme activity is not an immediate (minutes–hours) event, and is seen over a longer time span (hours–days) (Figure 2.12). Returning to baseline levels following inducing xenobiotic exposure also takes a similar time course as the "additional" enzymes undergo catabolism at the normal rate for the enzyme. Induction is not open-ended with regard to continued exposure; there are upper limitations to the extent or magnitude of induction. With therapeutic agents, induction may necessitate a change in the dosage regimen of drugs with time. With environmental xenobiotic exposures, also, induction may alter the biological effect from altered pharmacokinetics and changes in metabolic routes. As with inhibition, induction has been most extensively studied with cytochrome P450s although the same processes apply to some UGTs, GSTs, and ABC transporters. Enzymes and isozymes vary in their degree of inducibility. In general, the most inducible are CYPs (most), UGTs, and GSTs, while the most refractory are CYP2D6, FMOs, and SULTs.

Xenobiotic inducing agents generally have considerable lipid solubility and a relatively long biological half-life (i.e., they gain access to the liver and remain there for some finite period of time). While overly simplistic, induction can be viewed as a temporary adaptive response that occurs when the body (liver) is challenged with the possible accumulation of unwanted and possibly toxic xenobiotics (Figure 2.13). If the xenobiotic enters the liver cell and is adequately metabolized, and discharged as metabolites, there is little need for increasing the liver's metabolic capacity. If "excess" xenobiotic accumulates, either as a result of a high dose or because of continuous exposure, induction of its metabolism and secretion is a mechanism to reduce its concentration. With high levels of the xenobiotic, some is able to bind to a cytosolic receptor protein that has some affinity for the accumulating xenobiotic. The binding initiates a conformational change in the receptor, sometimes displacing a chaperone protein and is then able to translocate into the nucleus. In the nucleus, the receptor–agonist complex is able to link with a partner nuclear factor(s), often a retinoic acid binding protein, and the heterodimeric complex interacts with a regulatory (promoter) region of DNA (the interaction site is often designated as a DRE or XRE, a "drug" or "xenobiotic" response element). The receptors, nuclear factors, and response elements for the four major induction mechanisms (identified by their most characteristic inducing xenobiotic "agonist") that affect cytochrome P450 transcription are given in Table 2.10, with a listing in Table 2.11, and the structure of some prototype inducers are shown in Figures 2.14 and 2.15. This binding to DNA brings all the necessary proteins and enzymes together to initiate the transcription of mRNA of those genes that have the necessary DRE/XRE sequence, that is, there is some selectivity; not all xenobiotic-metabolizing enzymes will have the sequence.

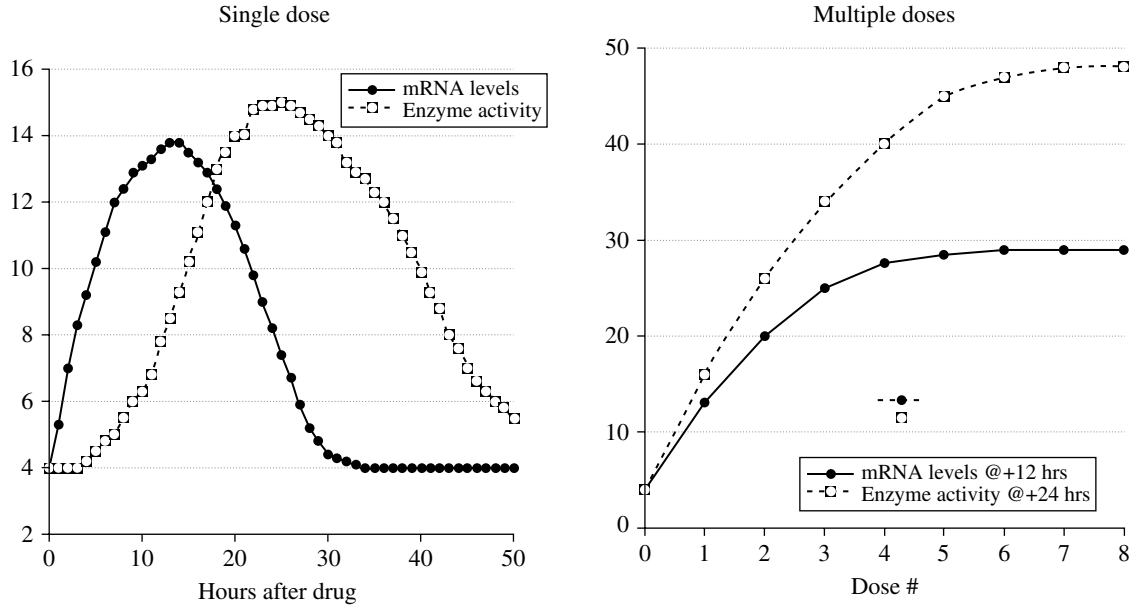


FIGURE 2.12 Activity and mRNA changes with inducer exposure.

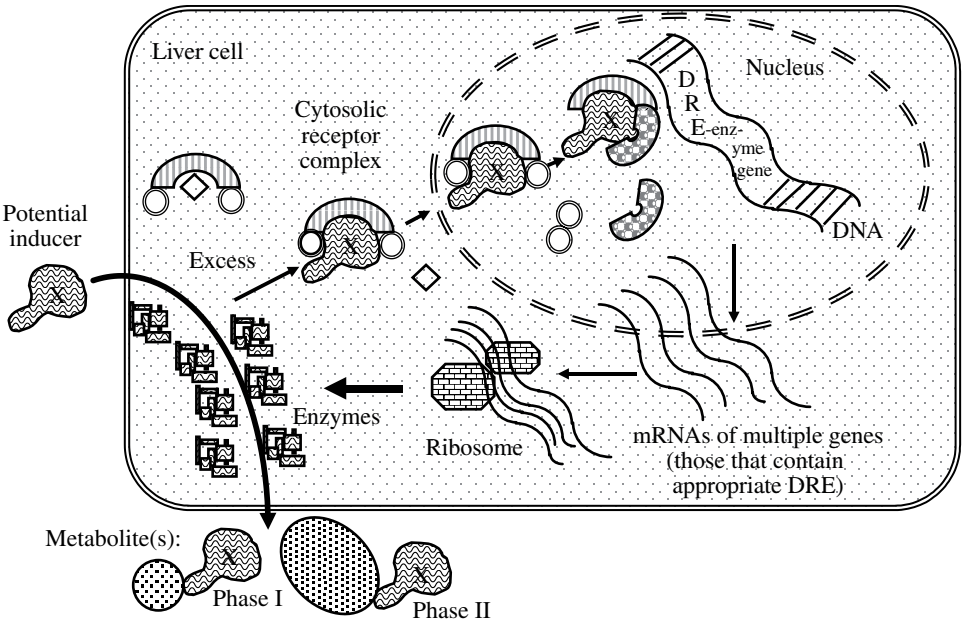


FIGURE 2.13 Xenobiotic-metabolizing enzyme induction.

TABLE 2.10 Components of Four Cytochrome P450 Induction Pathways

Pathway	1	2	3	4
Agonist	Rifampicin	Polycyclic aromatics TCDD	Phenobarbital	Clofibrate
	Hyperforin	Coplanar PCBs	Non-coplanar PCBs	Di(2-ethylhexyl-phthalate (DEHP)
Cytosolic Receptor	Pregnane X (PXR)[a]	Aromatic hydocarbon (AhR)[b]	Constitutive androstane (CARβ)	Peroxisome proliferator activated (PPAR)
Nuclear coregulator	9-cis retinoic acid receptor (RXR)	Ah Receptor nuclear translocator (Arnt)	9-cis retinoic acid receptor (RXR)	9-cis retinoic acid receptor (RXR)
DNA response Element	PXRE	DRE/XRE	PBRE	PPRE

[a]Ligands of the PXR receptor have a diverse array of structures: the binding site accommodates many shapes and sizes of molecules (see Figure 2.14).
[b]The AhR receptor has preference for large planar molecules (see Figure 2.15). Agonist binding releases the receptor from a sequestering complex comprising the AhR receptor itself, a 23 kDa cochaperone protein (p23), an X-associated protein (XAP2), and two 90-kDa heat shock proteins (hsp90).

TABLE 2.11 Cytochrome P450 Inducers

CYP	Xenobiotic Inducing Agent
1A2	Cruciferous vegetables[a] (broccoli, brussels sprouts), char-grilled meat (polycyclic aromatic hydrocarbons, e.g., benzopyrene), tobacco products, 3-methylcholanthrene, 1,2-benzanthracene TCDD, modafinil, nafcillin, β-naphthoflavone, omeprazole, rifampicin, lansoprazole, oltipraz
2A6	Phenobarbital, rifampicin, dexamethasone, pyrazole
2B6	Phenobarbital, rifampicin, dexamethasone, phenytoin
2C8/9	Phenobarbital, rifampicin, dexamethasone
2C19	Phenobarbital, rifampicin, dexamethasone, carbamazepeine, norethindrone, prednisone
2D6	None known[b]
2E1	Ethanol, isoniazid, pyridine, 4-methylpyrazole
3A4/5	Barbiturates, carbamazepine, ciglitazone, clotrimazole, dexamethasone, efavirenz, glucocorticoids, modafinil, nevirapine, oxcarbazepine, phenobarbital, phenytoin, pioglitazone, rifabutin, rifampicin, rifapentin, ritonavir, St. John's Wort [hyperforin](and some other nutraceuticals), sulfinpyrazone, taxol, troglitazone, troleandomycin

[a]Contain glucosinates that are metabolized to Ah receptor-activating indoles.
[b]Ironically, a CYP present in low abundance in an uninduced liver.

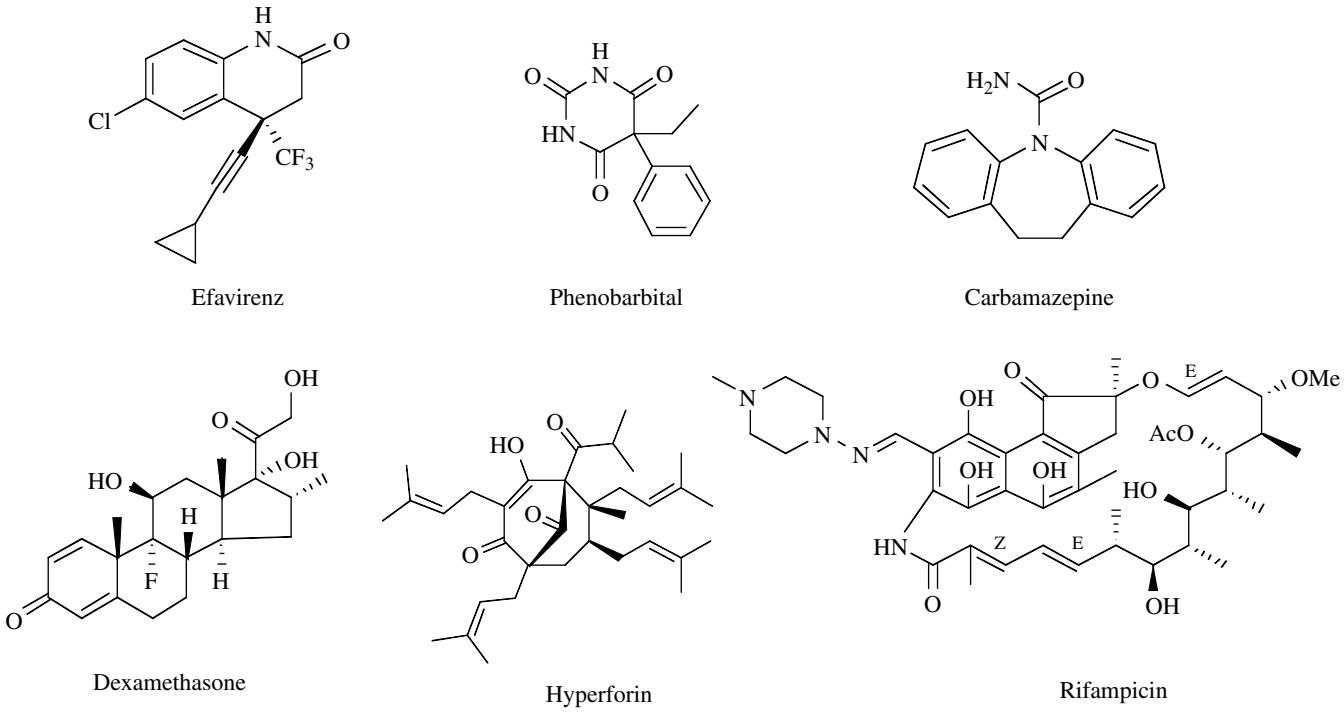


FIGURE 2.14 Examples of the diversity of pregnane X receptor (PXR) agonists.

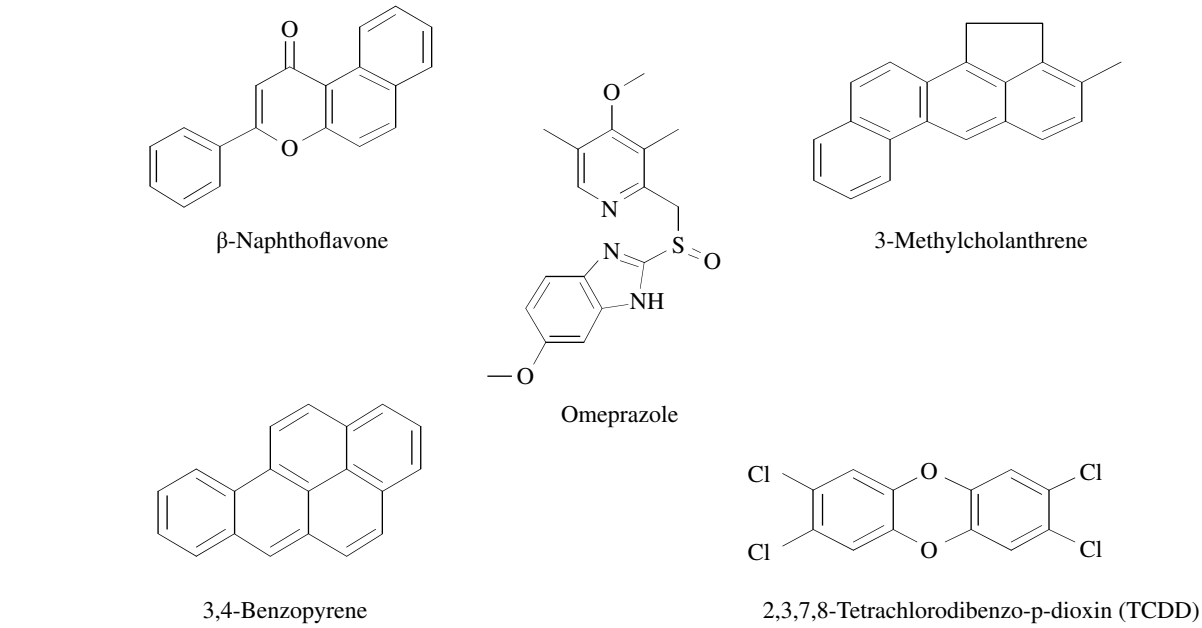


FIGURE 2.15 Examples of the planar nature of aromatic hydrocarbon receptor (AhR) agonists.

Enzyme and xenobiotic selectivity for induction therefore resides at the cytosolic receptor and the DRE/XRE (i.e., does the xenobiotic bind a suitable receptor and does the enzyme gene have the necessary recognition sequence for the occupied receptor complex). The mRNA molecules move out of the nucleus and are translated into new proteins (induced enzymes) on the ribosomes attached to the endoplasmic reticulum (Table 2.12).

There are five major "complexities" associated with induction.

1. A specific enzyme may be induced by a single mechanism by a range of xenobiotics. (The cytosolic receptor binds different molecules.)
2. A specific enzyme may be induced by multiple mechanisms, each with its associated range of xenobiotic

TABLE 2.12 Induction-Related Enhanced Drug Metabolism in Human

Inducer	Enhanced Metabolism
Benzo[a]pyrene	Theophylline
Carbamazepine	Carbamazepine, clonazepam, itraconazole
Chlorcyclizine	Steroid hormones
Ethchlorvynol	Warfarin
Glutethimide	Antipyrine, glutethimide, warfarin
Griseofulvin	Warfarin
Phenobarbital	Barbiturates, choramphenicol, chlorpromazine, cortisol, coumarin anticoagulants, desmethylimipramine, digitoxin, doxorubicin, estradiol, itraconazole, phenylbutazone, phenytoin, quinine, testosterone
Phenylbutazone	Aminopyrine, cortisol, digitoxin
Phenytoin	Cortisol, dexamethasone, digitoxin, itraconazole, theophylline
Rifampicin	Coumarin anticoagulants, digitoxin, glucocorticoids, itraconazole, methadone, metoprolol, oral contraceptives, prednisone, propranolol, quinidine, saquinavir
St John's Wort	Alprazolam, cyclosporine, digoxin, indinavir, oral contraceptives, ritonavir, simvastatin, tacrolimus, warfarin

agonists. (The enzyme may have more than one kind of response element.)

3. A single xenobiotic may be responsible for the induction of multiple enzymes (not necessarily to the same extent) by the same mechanism. (Each enzyme has the requisite response element.)
4. A single xenobiotic may induce multiple enzymes (not necessarily to the same extent) by different mechanisms. (The xenobiotic can bind to more than one kind of cytosolic receptor and thereby the enzyme can be induced by different mechanisms.) Induction by xenobiotics that occur as mixtures (e.g., polychlorinated biphenyls (PCBs)) are a variation of this in that the mixture may contain chemical entities that bind to different receptors.
5. Induction by one xenobiotic can affect the metabolism (and thereby the biological effects) of many others that are substrates for the induced enzyme(s).

The concept of the hepatocyte responding to the presence of "excess" xenobiotic is important in realizing that with concomitant exposure to multiple drugs or xenobiotics, some may act as inhibitors of P450 and cause others to accumulate above normally noninducing levels and become inducers.

In addition to the induction of CYPs, other enzymes and transporters can be induced by the same mechanisms. Coinduction of Phase II enzymes and transporters along with CYPs protects the induced liver cell from the accumulation of possibly toxic Phase I metabolites (Figure 2.5). Rifampicin has been shown to also induce UGT1A1 and UGT1A6, and MRP2 and MDR1; β-naphthoflavone also induces UGT1A1 and UGT1A6, and MRP2 and MDR1 (but via a different mechanism), and phenobarbital induces UGT1A1, and numerous transporters.

In addition to the induction mechanisms outlined in Table 2.10, another pathway exists that appears confined to "cytoprotective" or inactivating enzymes (i.e., those enzymes responsible for the detoxification or removal of reactive metabolites). This Kelch-like ECH associating protein 1 (Keap1)-nuclear factor E2-related protein 2 (Nrf2) pathway is initiated by the release of Nrf2 from its anchoring Keap1 tether (often mediated by oxidative or electrophilic stress) and in the nucleus, the Nrf2 binds to the regulatory regions (termed an antioxidant response element or ARE) of protective enzymes. Enzymes responding to this mechanism include NADPH–quinone oxidoreductase, numerous GSTs, and microsomal epoxide hydrolase. Activators of this Nrf2/ARE system include butylated hydroxyanisole (BHA) and butylated hydroxytoluene (BHT), epigallocatechin-3-gallate (a polyphenol from green tea), oltipraz, sulforaphane (an isothiocyanate from cuciferous vegetables such as broccoli), and various glutathione-depleting agents such as diethylmaleate.

In vivo, the consequences of induction can be in opposite directions. If metabolism of a xenobiotic results in inactivation, induction will likely decrease the biological (pharmacological or toxic) effect. In those instances where metabolism results in bioactivation of a drug or creation of a toxic/reactive metabolite, induction will likely increase the pharmacological effect and/or toxicity.

Whether a compound is a likely inducing agent *in vivo*, or not, is now often determined outside of whole-animal studies using short-duration *in vitro* cell-based systems. In some of the cell-based screening systems, the cells contain reporter enzymes that have been engineered to contain the response elements of the various xenobiotic-metabolizing enzymes.

2.5 EXCRETION (ELIMINATION)

Absorption, metabolism, and excretion are all interrelated (see Figure 2.2).

All occur within a short time period of each other, and metabolism and excretion occur continuously in synchrony thereafter.

Renal

Elimination via the kidneys is the major route of xenobiotic excretion. The elimination is mostly as metabolites, particularly Phase II conjugates. Three processes occur in the renal tubule, ultrafiltration, active transport, and diffusion, and net excretion is the combined outcome of these three processes. The first two of these processes occur sequentially along the nephron as glomerular filtration and tubular secretion. Diffusion can occur in any segment of the nephron where there is a chemical concentration gradient, but particularly in the distal tubule. Blood circulation covers the entire nephron, allowing all processes to occur in one pass (Figure 2.16).

Glomerular filtration is a passive process dependent on the pressure and flow of blood into the kidney. The glomerulus receives 10% of renal blood supply but filtration is an efficient process and so this makes a large contribution to elimination of xenobiotics and/or their metabolites. Only free xenobiotic or xenobiotic metabolite is filtered, and any that are protein (most often albumin)-bound are retained in the bloodstream since protein size exceeds the interstitial pore size.

Tubular secretion of organic xenobiotics occurs almost exclusively in the proximal tubule. It is a two-step process, where both steps, at least for Phase II metabolites, involve transporters (Figure 2.17). SLCs move the compound to be excreted from the bloodstream into the tubule cell and ABC transporters together with SLCs move them from the cell into the urine. As the name implies, ABC transporters utilize ATP as their energy source and the input of energy can allow for export from the cell to occur against a chemical concentration gradient. Serum protein binding of compounds does not have as dramatic an effect on secretion as it does on glomerular filtration because relative to the loose association with albumin, the transporters have greater or higher affinity and are able to strip drug or xenobiotic off the albumin. Multiple transport systems exist and they are variable with the nephron region; cation/basic carriers occur in both the early section and the midsection of the proximal tubule while anion/acid carriers are concentrated in the midsection. General transporter characteristics apply to both SLCs and ABCs; they exhibit selectivity, and can experience competition and saturation. The ABCs comprise MRPs (products of *ABCC* genes) and MDRs (products of *ABCB*) genes. SLCs comprise organic cation transporters (OCTs) and organic anion transporters (OATs). SLCs differ from ABCs in having no ATP binding site and rely more on linked cotransport (secondary active transport) to drive the movement. Structurally, SLCs and ABCs are very similar, each having upward of 12 membrane-spanning regions that are able to orient to form a "cylindrical" pore or channel through which the transported molecule moves. Drugs subject to active anion transport include probenecid, penicillins, cephalosporins and sulfonamides, NSAIDs, loop diuretics such as furosemide and ethacrynic acid, thiazide diuretics, and, deserving of special recognition, ionized Phase II metabolism conjugates, notably glucuronides and sulfates. Compounds subject to active cation transport include amiloride, neostigmine (a quaternary ammonium drug), quinine, histamine, H2 antagonists, morphine, ephedrine and pseudoephedrine, choline, and thiamine.

Countering the active secretion process in the proximal tubule is back diffusion from the urine into the bloodstream through the lipid bilayer cell membrane. The efficiency of this process obviously depends on the lipid solubility of the compound (lipid/water partition coefficient), a parameter that is higher for unchanged xenobiotic that was originally absorbed by diffusion in the GI tract than for metabolites, especially

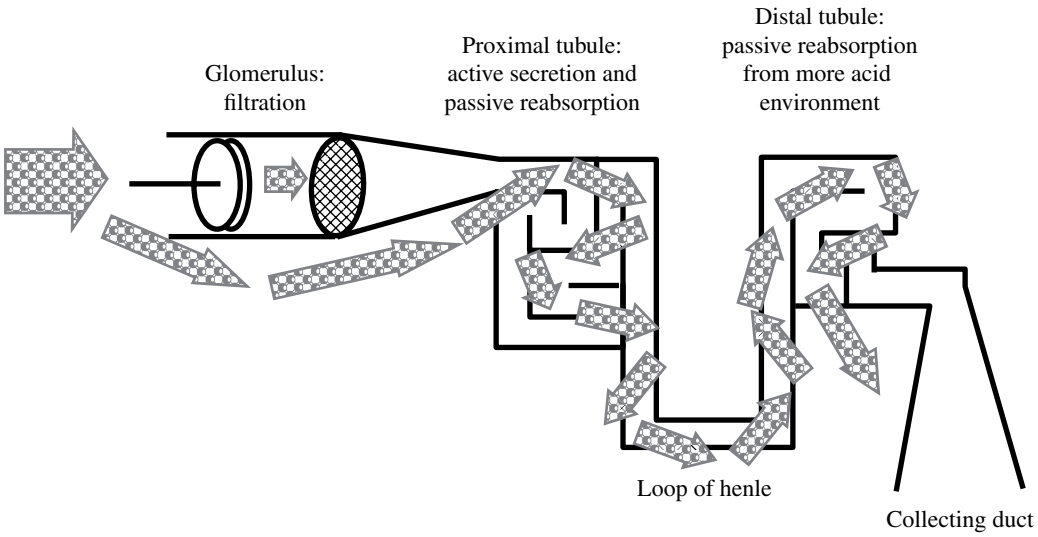


FIGURE 2.16 Processes linked to the renal elimination of xenobiotics.

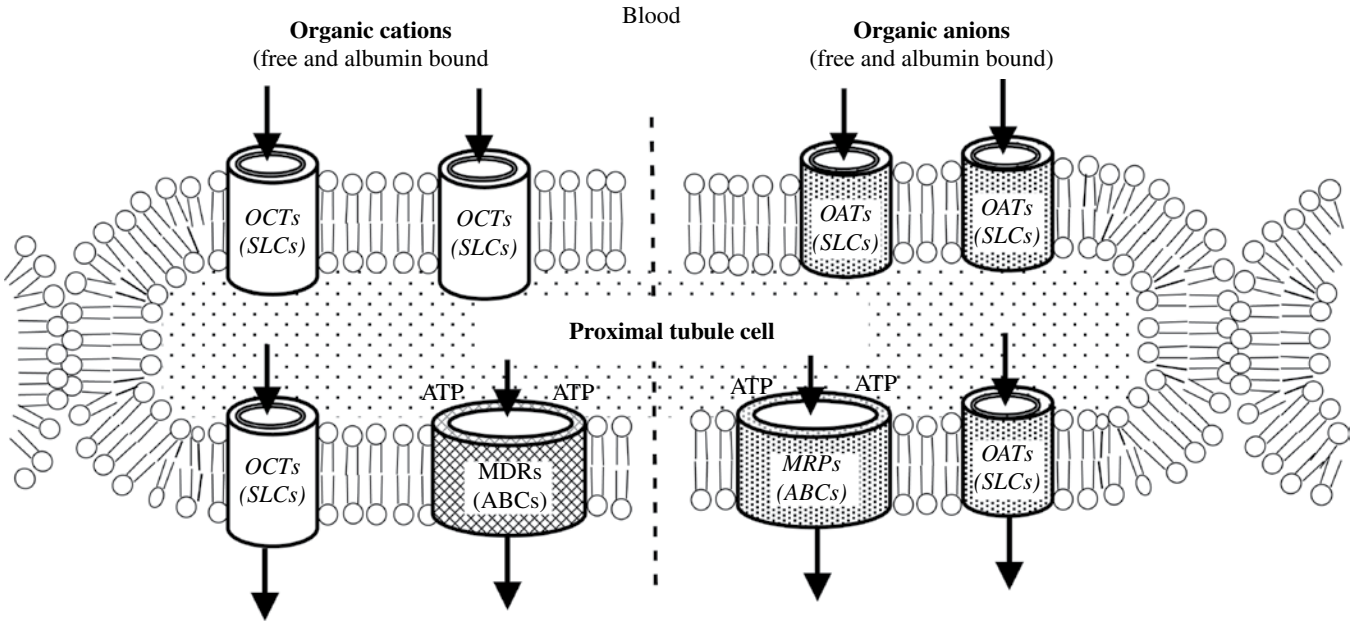


FIGURE 2.17 Transporter involvement in renal xenobiotic excretion.

Phase II metabolites existing extensively in the ionized state at physiological pHs. Through the exchange of ions as the urine moves through the nephron, it becomes slightly more acidic relative to the original glomerular exudates. The acidic pH will decrease the degree of ionization of weak acids and enhance their back diffusion (reabsorption). For basic compounds, which for xenobiotics is largely an amine moiety, the acidic pH will increase the extent of ionization and decrease the tendency for reabsorption. This pH (ion trapping) effect can and has been utilized to enhance the excretion of drugs from the body. Alkalinization of the urine (e.g., from intravenous sodium bicarbonate administration) will enhance the excretion rate of anions and has been used in barbiturate overdose situations. Other weak acids showing increased excretion in alkaline urine include chlorpropamide, methotrexate, salicylates, and sulfonamides. Acidification of the urine (e.g., high ascorbic acid intake) will enhance the clearance of amine drugs such as amphetamine from the body. Examples of weak bases showing increased excretion in acidic urine include ephedrine, pseudoephedrine, mexilitine, and quinine.

The net outcome of the three kidney processes amounts to a "renal extraction ratio" that indicates the difference between the concentration of xenobiotic in the blood entering the kidney and the amount in the blood leaving. Drugs such as penicillin and Phase II conjugates such as glucuronides and sulfates have high extraction ratios (>0.7). Drugs with low extraction ratios (<0.3) include atenolol, cefazolin, chlorpropamide, digoxin, furosemide, gentamicin, phenobarbital, sulfisoxazole, and tetracycline, and such compounds require several passes to achieve sizable excretion.

Kidney (renal) excretion is not the only xenobiotic elimination pathway. Among others are the GI tract, lungs, and breast milk.

Biliary Excretion

Xenobiotics that enter the hepatocyte from the bloodstream (much of which has come from the intestinal site of absorption via the portal vein), either by diffusion or by way of OATs and OCTs, can return to the bloodstream (exiting the liver lobule via the central vein) by diffusion (not likely if lipid-soluble xenobiotic entering has been converted to a more water-soluble metabolite) or ABC transporter–mediated processes. Returning to the bloodstream will make the xenobiotic and/or its metabolite(s) available for subsequent renal excretion (Figure 2.16). An alternative exit pathway from the liver is via the bile (Figure 2.2). ABC transporters on the canalicular membrane mediate this process. The bile eventually enters the GI tract at the proximal end of the small intestine and as long as the xenobiotic or its metabolite has a low lipid/water partition coefficient (as is likely if it has been metabolized to an extensively ionized conjugate in the liver), it will not undergo absorption and will pass out in the stool. For some metabolites, there are enzymes present in the GI tract that can "deconjugate" metabolites (e.g., β-glucuronidases) resulting in a return of the xenobiotic or its Phase I

metabolite that likely has greater lipid solubility and can be reabsorbed. If a xenobiotic undergoes reabsorption and is returned to the liver for "remetabolism" it is said to undergo enterohepatic circulation. Fortunately, enzymes such as the β-glucuronidases largely arise from resident bacteria in the lower regions of the GI tract, distal to the large absorptive surface of the small intestine.

Respiration

Exhalation of xenobiotics is only an efficient elimination pathway for gases or compounds with a high vapor pressure. In terms of pharmacological agents, it is the major excretory pathway for the gaseous general anesthetics following surgery. For environmental and workplace chemicals present as gases, and vapors, the lungs are both a portal of entry and a site of elimination. For the socially acceptable drug, ethanol, equilibration between blood and alveolar air is the basis of the breathalyzer test for determining blood alcohol levels, but in terms of overall ethanol elimination, exhalation of unmetabolized ethanol is an extremely minor component.

Milk

Although only a sporadic route by which xenobiotics and xenobiotic metabolites leave the body in humans, its occurrence in dairy animal species can be of consequence to downstream consumers. The concentration in milk depends upon the concentration of the free xenobiotic in plasma and on physicochemical properties of the xenobiotic, being of greatest concern for lipid-soluble compounds. In humans, xenobiotics in maternal milk can result in chronic xenobiotic exposure to infants at a time when metabolism and elimination mechanisms are not fully developed. Indeed, warnings have been issued following an infant mortality from a morphine overdose in a nursing mother taking codeine.

2.6 SUMMARY

- Lipophilicity and hydrophilicity are key determinants in the absorption, distribution, and excretion of xenobiotics.
- Metabolism of xenobiotics occurs in one or two phases; each phase generally bestows increasing hydrophilicity and excretability to the molecule.
- Metabolism proceeds by many different chemical mechanisms and is catalyzed by many different classes of enzymes. The enzymes exhibit substrate (xenobiotic) selectivity.
- Metabolism may increase or decrease the biological reactivity (toxicity) of a xenobiotic.
- Xenobiotic metabolism can be highly variable from influences of external factors and genetics.

SUGGESTED READING

David Rodrigues A, editor. *Drug-Drug Interactions.* 2nd ed. New York: Informa Healthcare; 2008. Various Chapters.

Hodgson E, editor. *A Textbook of Modern Toxicology*. 4th ed. Hoboken: John Wiley & Sons, Inc.; 2010. Chapters 6–10.

Ioannides C, editor. *Cytochromes P450: Role in the Metabolism and Toxicity of Drugs and other Xenobiotics*. Cambridge: Royal Society of Chemistry Publishing; 2008. Various Chapters.

Klaassen CD, editor. *Casarett and Doull's Toxicology: The Basic Science of Poisons*. 7th ed. New York: McGraw-Hill; 2008.

Lehman-McKeeman LD. Absorption, distribution, and excretion of toxicants. In: Klaassen CD, editor. *Casarett and Doull's Toxicology: The Basic Science of Poisons*. 7th ed. New York: McGraw-Hill; 2008.

Parkinson A, Brian WO. Biotransformation of xenobiotics. In: Klaassen CD, editor. *Casarett and Doull's Toxicology: The Basic Science of Poisons*. 8th ed. New York: McGraw-Hill; 2008.

Timbrell JA, editor. *Principles of Biochemical Toxicology.* 4th ed. New York/London: Informa Healthcare; 2008. Chapters 3–5.

3

TOXICOKINETICS

Rebecca A. Clewell and Harvey J. Clewell III

The intent of this chapter is to provide a general overview of the concepts and methods associated with characterizing the absorption, distribution, metabolism, and elimination of toxic chemicals in the body. The concepts described in this chapter are key elements for understanding the mode of action for the toxicity of a chemical as well as for estimating the risks associated with exposures to the chemical. This chapter explains:

- The factors that determine the toxicokinetic (TK) behavior of a chemical
- The methods that are applied to investigate and describe TK behaviors
- Common applications of these methods in toxicology and risk assessment

3.1 INTRODUCTION

TK is the quantitative study of factors that control the time course for absorption, distribution, metabolism, and excretion (ADME) of toxic compounds within the body. The time course of drugs, on the other hand, has traditionally been referred to as pharmacokinetics (PK). Similarly, the effect of a drug or toxic compound in the target tissue for its effect has been referred to as pharmacodynamics (PD) and toxicodynamics (TD). This practice, of course, ignores the wisdom of Paracelsus: "Only the dose differentiates a poison and a remedy." To avoid this false distinction, the terms biokinetic and biodynamic have sometimes been used. In this chapter these alternative terms will be used interchangeably. Biokinetic models provide sets of equations that simulate the time courses of compounds and their metabolites in various tissues throughout the body.

The interest in biokinetics in pharmacology and toxicology arose from the need to relate internal concentrations of active compounds at their target sites to the doses of the compound administered to an animal or human subject. The reason for this interest is a fundamental tenet in pharmacology and toxicology—that both beneficial and adverse responses to compounds are related to the free concentrations of active compounds reaching target tissues, rather than the amount of compound at the site of absorption. The relationship between tissue dose and administered dose can be complex, especially at high doses, with repeated daily dosing, or when metabolism or toxicity at routes of entry alter uptake processes for various routes of exposure. Biokinetic models of all kinds are primarily tools to assess internal dosimetry in target tissues for a wide range of exposure situations.

Absorption, Distribution, Metabolism, and Elimination

Environmental chemicals cover a wide range of physicochemical properties from lipophilic ("fat loving") to hydrophilic ("water loving"), highly volatile to nonvolatile, soluble to insoluble, and highly reactive to inert. These physiochemical properties determine whether the compounds are taken up into the body by various routes (inhalation, oral absorption, dermal absorption, etc.), whether they will be subject to metabolic activation or deactivation and whether they will be quickly or slowly cleared from the body. As a result, it is difficult to predict their TK behavior. Nevertheless, some general statements can be made.

Principles of Toxicology: Environmental and Industrial Applications, Third Edition. Edited by Stephen M. Roberts, Robert C. James and Phillip L. Williams.

Absorption For chemicals in the environment, the principal routes of exposure are ingestion, inhalation, and dermal contact. In general, environmental chemicals are well absorbed from either the gastrointestinal (GI) tract or the alveoli in the lung. Important exceptions to this rule include metals and insoluble particulates, which are often poorly bioavailable. The skin, on the other hand, is designed to serve as a barrier, and the bioavailability of chemicals is typically much lower by this route. Experimental studies in animals by artificial means such as intravenous or intraperitoneal dosing must be carefully evaluated, considering the differences in the concentration profiles for a chemical administered internally rather than by a physiologically relevant route.

Distribution Once absorbed, most environmental compounds are distributed throughout the body in the blood and are rapidly taken up into the various organs and tissues. Importantly, physiological "barriers" such as the placenta and the blood–brain barrier are often ineffective at restricting the movement of environmental contaminants.

Metabolism The same types of enzymes that metabolize endogenous chemicals and drugs are also capable of metabolizing environmental compounds. These include "Phase 1" enzymes, such as the cytochrome P450 oxidases (cyps) and esterases, as well as "Phase 2" enzymes that conjugate chemicals with glutathione, glucuronic acid or sulfate, making them more water soluble and more readily excreted in the urine. While metabolism often serves to detoxify an active compound, some metabolism, particularly by cyps and glutathione conjugation, can make a chemical more toxic.

Elimination The primary routes of elimination are exhalation (for volatile chemicals), urinary excretion (for water-soluble compounds), and fecal elimination. Elimination may be either passive (e.g., glomerular filtration of free chemical in the blood) or active (e.g., biliary or renal secretion).

Toxicokinetic Modeling

In the 1930s, Teorell provided a set of equations for uptake, distribution, and elimination of drugs from the body. These equations are rightly regarded as providing the first physiological model for drug distribution. However, computational methods were not available to solve the sets of equations at that time. Exact mathematical solutions for distribution of compounds in the body could only be obtained for simplified models in which the body was reduced to a small number of compartments that did not correspond directly with specific physiological compartments. Over the next 30 years, modeling of drugs focused on these simpler descriptions with exact solutions rather than on developing models more concordant with the structure and content of the biological system itself. These approaches are sometimes referred to as "data-based" compartmental modeling since the work generally took the form of a detailed collection of time-course blood/excreta concentrations at various doses. Time-course curves were analyzed by assuming particular model structures and estimating a small number of model parameters by curve fitting. In the earliest of these models, all processes for metabolism, distribution, and elimination were treated as first-order rates (i.e., rates changed in direct proportion to the concentration of the chemical species). In the 1960s and early 1970s concerns were raised about the ability of data-based compartmental modeling to account for (i) the saturation of elimination pathways and (ii) the possibility that blood flow, rather than metabolic capacity of an organ, might limit the clearance of a chemical. Saturation of elimination led to models that were not first-order, making it difficult to derive exact solutions to the sets of equations, while blood flow–limited metabolism in an organ meant that the removal rate constant from a central compartment could not increase indefinitely as the metabolic capacity increased.

Scientists trained in chemical engineering and computational methods developed more complicated, physiologically based pharmacokinetic (PBPK) models for chemotherapeutic compounds. Many of these compounds are highly toxic and have therapeutic efficacy by being slightly more toxic to rapidly growing cells (i.e., cancer cells) than to normal tissues. The more complete physiological structure of the models allowed the investigators to determine the impact of different dosing scenarios to maximize efficacy in the tumor while minimizing toxicity to other tissues. These seminal contributions showed the ease with which realistic descriptions of physiology and relevant pathways of metabolism could be incorporated into PBPK models and paved the way for their more extensive use for both pharmaceutical and environmental compounds.

Today, there are three major types of computation models used for the analysis of biokinetic data: noncompartmental analysis (NCA), classical compartmental models similar to those described earlier, and PBPK models. NCA is widely used in the pharmaceutical industry for the analysis of single-dose PK data and PK data generated during safety studies. NCA, which involves calculating parameters such as volume of distribution (V_{ss}), clearance (CL), area under the plasma concentration curve (AUC), and peak plasma concentration (C_{max}), based solely on experimental data (i.e., these calculations are model-independent), is a convenient way to understand, tabulate, and compare the PK properties of compounds. However, NCA results cannot be used to extrapolate to other exposure conditions.

Classical compartmental models can be used to perform simulations (e.g., use data generated from one exposure

condition to simulate a different exposure condition). For example, a compartmental model parameterized with data from a single exposure can be used to predict the expected behavior for repeated exposure, assuming there is no change in the TK of the chemical over the period of the exposure. Standard software packages often provide flexible tools for developing one-, two-, and three-compartmental models from time-course plasma concentrations, and provide statistical tools for discriminating between models. However, the parameters and compartments in compartmental models have no physiological meaning and provide no basis for extrapolating from the conditions of the experimental data used to develop the model (e.g., to a different exposure route or a different species).

PBPK models differ from classical compartmental models in that they include specific compartments for tissues involved in exposure, toxicity, biotransformation, and clearance processes. Tissues are connected by blood flow and compartments and blood flows are described using physiologically meaningful parameters. Figure 3.1 illustrates the structure of a simple PBPK model for a volatile, lipophilic compound—styrene. As with compartmental models, PBPK models can be used to simulate a variety of conditions. However, because PBPK models utilize realistic parameters for tissue volumes and kinetic processes, they can be used for extrapolation across doses, exposure routes, and even species. PBPK modeling has been used to great effect for interspecies extrapolation, both among animal models and for predicting human kinetics based on animal data. The mechanistic basis of PBPK models allows for applications such as understanding species differences in target tissue chemical exposure, determining if results from different experimental designs are consistent, and exploring possible mechanisms responsible for unexpected or unusual data. These attributes have led to widespread development of PBPK models in recent years.

PBPK models have many advantages over the more empirical descriptions of chemical kinetics (NCA, classical compartmental model), perhaps the most important of which is their greater predictive power. Since known physiological parameters are used, a different species can be modeled by simply replacing the appropriate constants with those for the species of interest or by allometric scaling (scaling proportional to body weight raised to an empirically determined power). Similarly, the behavior for a different route of administration can be determined by adding equations that describe the nature of the new input function.

Another important benefit is the reduced need for extensive experiments with new compounds. Since measured

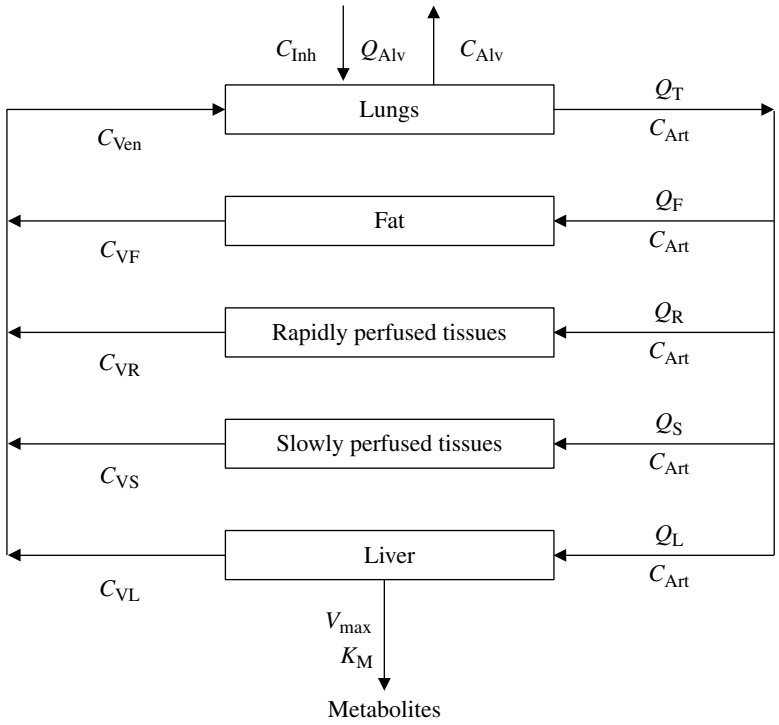


FIGURE 3.1 Diagram of a physiologically based pharmacokinetic (PBPK) model for styrene. In this description, groups of tissues are defined with respect to their volumes, blood flows (Q), and partition coefficients for the chemical. The uptake of vapor is determined by the alveolar ventilation (Q_{ALV}), cardiac output (Q_T), blood:air partition coefficient, and the concentration gradient between arterial and venous pulmonary blood (C_{ART} and C_{VEN}). Metabolism is described in the liver with a saturable pathway defined by a maximum velocity (V_{max}) and affinity (K_m). The mathematical description assumes equilibration between arterial blood and alveolar air as well as between each of the tissues and the venous blood exiting from that tissue. *Source:* Adapted from Ramsey and Andersen (1984).

physical-chemical and biochemical parameters are used, the behavior for new compounds can quickly be estimated by determining the appropriate constants for key processes, such as metabolism. The process of selecting the most informative experimental data is also facilitated by a predictive PBPK model. Fundamentally, physiologically based models provide a conceptual framework for employing the scientific method: hypotheses can be described in terms of biological processes, quantitative predictions can be made on the basis of the mathematical description, and the model (hypothesis) can be revised on the basis of comparison with targeted experimental data.

The trade-off for the greater predictive capability of physiologically based models is the requirement for an increased number of parameters and equations compared to the more empirical models. However, values for many parameters, particularly physiological parameters, are readily available in the literature. Furthermore, many *in vitro* techniques have been developed for rapid determination of compound-specific parameters, such as those describing tissue partitioning and metabolism. An important advantage of PBPK models is that they provide a biologically meaningful quantitative framework wherein *in vitro* data can be more effectively utilized to predict *in vivo* TK. There is even a prospect that predictive PBPK models can be developed based almost entirely on data obtained from *in vitro* studies and quantitative structure–activity relationship (QSAR) modeling, eliminating the need for the use of animals in TK analyses. This effort to eliminate the need for animal-based kinetic studies through the use of PBPK models, QSAR modeling, and *in vitro* data collection is generally referred to as "*in vitro* to *in vivo* extrapolation," or IVIVE.

PBPK models are often developed in support of chemical risk assessments. A properly validated PBPK model can be used to perform the high- to low-dose, dose-route, and interspecies extrapolations that are necessary for estimating human risk from toxicology studies in laboratory animals (e.g., *in vivo* hazard studies). PBPK models are also used for examining the effects of changing physiology on target tissue dosimetry, as in the case of early life exposure where rapid organ growth and development of important metabolic enzymes affect chemical kinetics. Target tissue dosimetry provided by PBPK modeling is also an essential component in models of chemical-induced tissue effects (PD), such as acetylcholinesterase inhibition.

3.2 TOXICOKINETIC MODELING FUNDAMENTALS

Kinetic Processes

The rate of change in mass of chemical per unit of time (dA/dt) for a kinetic process is described in terms of concentration (C) rather than amount (A), since the free concentration of a chemical is the driving force for its kinetics. The "order" of the process refers to the power to which concentration is raised in the kinetic equation. The qualities of zero-, first-, and second-order rates are described here.

- *Zero-order process*
  - Rate is independent of concentration: ($dA/dt = k_0 \times C^0 = k_0$).
  - Rate constant (k_0) has units of mass per time (e.g., mg/h).
  - A plot of concentration over time is a straight line on a linear plot (Figure 3.2).
- *First-order process*
  - Rate is proportional to concentration: ($dA/dt = k_1 \times A = k_1 \times C \times V$),where V is the volume of the compartment with concentration C.
  - Rate constant (k_1) has units of reciprocal time (e.g., l/h).
  - A plot of concentration over time is a straight line on a logarithmic plot (Figure 3.3).

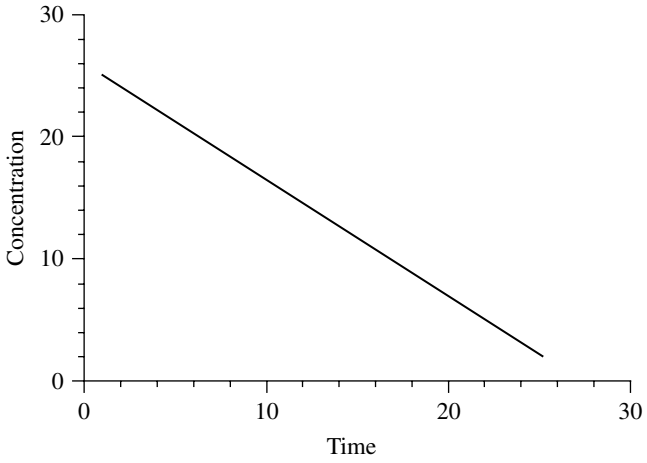


FIGURE 3.2 Zero-order reaction.

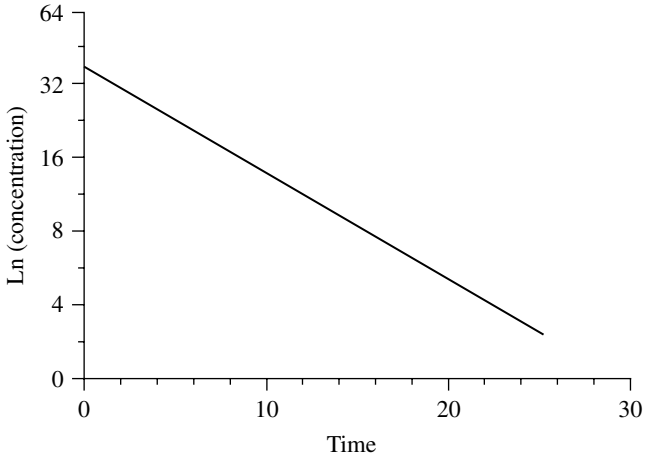


FIGURE 3.3 First-order reaction.

  - The half-life ($t_{1/2}$) of a chemical with first-order elimination can be defined as the time it will take for the concentration to decrease to half of its current value.
- *Second-order process*
  - Rate is proportional to two concentrations ($dA/dt = k_2 \times C_1 \times C_2 \times V$).
  - Rate constant (k_2) has units of per time per concentration ($t^{-1}C^{-1}$).
    - (e.g., l/mg/h).

An example of a second-order biokinetic process is the conjugation of a chemical with glutathione, which depends on both the concentration of the chemical and the concentration of glutathione. Under typical physiological conditions the concentration of glutathione is relatively large (several millimolar), such that the effect of the conjugation reaction on its concentration is negligible. The rate is then referred to as "pseudo-first order," since it is approximately proportional to the concentration of the chemical alone ($dA/dt = k_2 \times C_1 \times C_2 \times V \approx k_1 \times C_2 \times V$, where $k_1 = k_2 \times C_1$). However, if the conjugation reaction results in a decrease in the concentration of glutathione, the true second-order rate equation must be used.

- *Saturable (capacity-limited) process*
  - First-order at low concentrations, transition to zero-order at high concentrations (Figure 3.4).
  - Described using the Michaelis–Menten (M–M) equation:

$$\text{Rate} = (V_{\max} \times C) / (C + K_m)$$

where $V_{\max}$ (the enzyme capacity) is the maximum rate (amount per time); and K_m (the enzyme affinity) is the concentration (amount per volume) at which the rate is half $V_{\max}$.

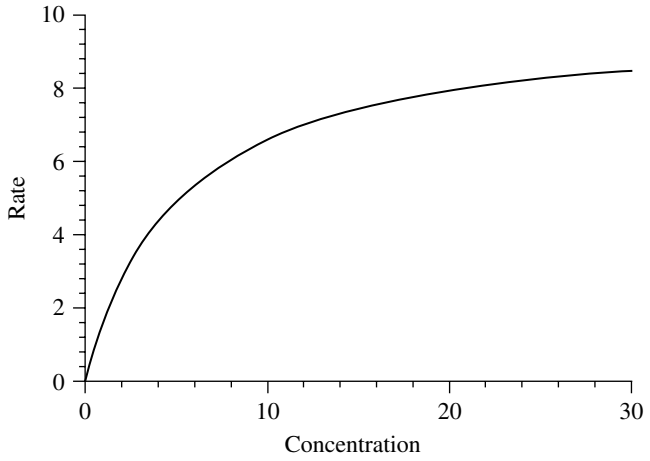


FIGURE 3.4 Capacity-limited reaction.

Compartmental Analysis

Classical compartmental modeling is largely an empirical exercise, where data on the time course of the chemical of interest in blood (and perhaps other tissues) are collected. Based on the behavior of the data, a mathematical model is selected, which possesses a sufficient number of compartments (and therefore parameters) to describe the data. The compartments do not generally correspond to identifiable physiological entities but rather are abstract concepts with meaning only in terms of a particular calculation. In a one-compartmental model, all of the compound in the tissues is assumed to be in equilibrium with the blood. If there is evidence of a storage depot for the chemical (e.g., the fat for a lipophilic compound), a second compartment can be added to the model and rate constants for transport between the central and deep compartments can be estimated from the data.

Using a one-compartment description, the change in serum concentration (C_s, mg/l) resulting from a given dose rate (D, mg/kg/day) is:

$$\frac{dC_s}{dt} = D - k \times V_d \times C_s$$

where V_d = apparent volume of distribution (fraction of body weight; l/kg); and k = rate constant for elimination (day^{-1}) = $0.693/t_{1/2}$ where 0.693 is the natural logarithm of 2, the factor by which the concentration changes over one half-life.

For one-compartment kinetics, the concentration of a compound at steady state (C_{ss}) is related to the daily intake (D) by the clearance of the compound, which in turn is related to the half-life ($t_{1/2}$) and volume of distribution (V_d). At steady state,

$$\frac{dC_s}{dt} = 0,$$
$$\text{so } C_{ss} = D / (k \times V_d),$$
$$\text{or } D = C_{ss} \times (0.693 \times V_d / t_{1/2}).$$

Therefore, to convert a blood concentration to an intake rate, both the half-life and the volume of distribution must be experimentally determined.

Clearance Modeling

The kidney removes chemicals from circulating blood by filtration. The liver removes chemicals from the circulating blood by metabolism. With each of these organs, we can describe the function of the organ in terms of "clearances." In this usage, clearance is a volumetric flow of blood (for instance, liters/hour) from which all chemical is removed. All clearances are expressed as liters/hour or as an equivalent

flow term. For any particular compartment we can speak of clearance by several different processes, including:

- Flow clearances
- Diffusional clearances
- Metabolic clearances
- Excretory clearances

Renal (Kidney) Clearance For the kidney, urinary clearance ($\mathrm{Cl}_{\mathrm{urine}}$) is estimated by the ratio of the total amount of chemical excreted in the urine over a given time interval divided by the blood concentration and duration of collection.

$$\mathrm{Cl}_{\mathrm{urine}} = \frac{\left(\left[C_{\mathrm{urine}} \times \text{urine volume}\right] / C_{\mathrm{blood}}\right)}{\text{collection duration}}$$

Thus, urinary clearance becomes the volumetric flow of blood from which the chemical would have to be completely removed to account for the observed excretion into the urine.

Another useful concept is extraction, that is, the proportion of blood flow from which all chemical is removed during a single pass through the organ. From the example with the kidney, extraction can be related to clearance and blood flow to the kidney (Q_{kidney}).

$$\text{Extraction}\left(E\right) = \frac{\mathrm{Cl}_{\mathrm{urine}}}{Q_{\mathrm{kidney}}}$$

Hepatic (Liver) Clearance A great deal of work and analysis has been conducted to describe the removal of drugs and toxicants by metabolism in the liver in relation to extraction and clearance. The major relationships are similar to those for the kidney.

$$\begin{aligned}\mathrm{Cl}_{\mathrm{liver}} &= Q_{\mathrm{liver}} \times \text{extraction} = Q_{\mathrm{liver}} \times \left(C_{\mathrm{ART}} - C_{\mathrm{VEN}}\right) / C_{\mathrm{ART}} \\ &= Q_{\mathrm{liver}} \times \left(C_{\mathrm{IN}} - C_{\mathrm{OUT}}\right) / C_{\mathrm{IN}}\end{aligned}$$

For the case where extraction is due to metabolism, the clearance at low substrate concentrations is readily expressed in relation to liver blood flow (Q_{liver}) and the kinetic parameters for metabolism, $V_{\max}$ and K_{m}.

$$\mathrm{Cl}_{\mathrm{liver}} = \frac{\left(Q_{\mathrm{liver}} \times V_{\max} / K_{\mathrm{m}}\right)}{\left(Q_{\mathrm{liver}} + V_{\max} / K_{\mathrm{m}}\right)}$$

As with the kidney, the interpretation for liver clearance is the volumetric flow of blood from which the chemical would have to be completely removed to account for the extraction. In a PBPK model the parameters in this equation correspond to measurable biological and biochemical entities. Q_{liver} is the actual liver blood flow (l/h), $V_{\max}$ is the maximum capacity of the metabolizing enzyme (l/h), and K_{m} is the affinity (mg/l). Both of the kinetic parameters can be estimated directly by *in vitro* experiments.

In deriving these parameters from compartmental PK models, we can also talk about clearance of compounds from the central compartment by the liver metabolism, although the parameters no longer correspond to their biological or biochemical counterparts. A one-compartmental model with metabolic elimination in the liver is expressed in terms of the amount in the compartment (A_1), the volume of distribution ($V_{\mathrm{d,1}}$), and the concentration in the compartment (C_1). The mass balance equation for the change in amount in the compartment can be written in several equivalent forms.

$$\frac{dA_1}{dt} = -k_{\mathrm{elim}} \times A_1$$

$$\frac{dA_1}{dt} = -k_{\mathrm{elim}} \times V_{\mathrm{d,1}} \times C_1$$

$$\frac{dA_1}{dt} = -\mathrm{Cl}_{\mathrm{liver}} \times C_1$$

In the last formulation, the loss of chemical from the system over time is liver clearance multiplied by the concentration in the central compartment. If there are other organs that are involved in removal, that is, in filtration by the kidney or in exhalation by the lungs, the equation is simply altered to account for the sum of all the clearances.

$$\frac{dA_1}{dt} = -\left(\mathrm{Cl}_{\mathrm{liver}} + \mathrm{Cl}_{\mathrm{kidney}} + \mathrm{Cl}_{\mathrm{exhalation}}\right) \times C_1$$

Restrictive and Nonrestrictive Hepatic Clearance By far, most of the metabolism of a compound occurs in the liver. In the standard description of clearance of a drug compound from blood by liver metabolism, binding in the blood is assumed to be linear and the fraction unbound (f_{ub}), which is assumed to be the fraction available for metabolism, is simply multiplied by the intrinsic hepatic clearance, $\mathrm{Cl}_{\mathrm{int}}$, leading to a straightforward relationship:

$$C_{\mathrm{L}} = \frac{Q_{\mathrm{L}} \times f_{\mathrm{ub}} \times \mathrm{Cl}_{\mathrm{int}}}{Q_{\mathrm{L}} + f_{\mathrm{ub}} \times \mathrm{Cl}_{\mathrm{int}}}$$

In this relationship, the maximal hepatic clearance, even with a low fraction unbound, is total tissue blood flow, Q_{L}. That is, all of a compound in the blood, whether bound or free, becomes available for clearance as long as the intrinsic

clearance is sufficiently large. Note that in this description the fraction unbound is not a function of the clearance. If the uptake of a compound into the metabolizing tissue is limited by the rate of dissociation of the compound from binding proteins in the blood or the rate of uptake into the tissue (referred to as restrictive clearance), the simple formula given earlier may overestimate its clearance. In the limit, the clearance for an irreversibly bound chemical would be:

$$C_{\mathrm{L}} = \frac{Q_{\mathrm{L}} \times f_{\mathrm{ub}} \times \mathrm{Cl}_{\mathrm{int}}}{Q_{\mathrm{L}} + \mathrm{Cl}_{\mathrm{int}}}$$

On the other hand, if dissociation and hepatocellular uptake are fast compared to the rate of clearance, the limiting behavior (nonrestrictive clearance) is given by:

$$C_{\mathrm{L}} = \frac{Q_{\mathrm{L}} \times \mathrm{Cl}_{\mathrm{int}}}{Q_{\mathrm{L}} + \mathrm{Cl}_{\mathrm{int}}}$$

Most modeling of environmental chemicals typically assumes that the clearance is nonrestrictive; that is, all of the compound in the blood entering the liver is assumed to be available for metabolism. In the case of oral exposure, the concentration of a chemical is determined by the ratio of the dose rate (mg/kg/day) to the sum of all clearances. For a chemical that is cleared by hepatic metabolism and urinary excretion, the resulting equation would be:

$$C_{\mathrm{ss}} = \frac{k_0}{\left[\mathrm{GFR} \times f_{\mathrm{ub}} + Q_{\mathrm{L}} \times f_{\mathrm{ub}} \times \mathrm{Cl}_{\mathrm{int}} / \left(Q_{\mathrm{L}} + f_{\mathrm{ub}} \times \mathrm{Cl}_{\mathrm{int}}\right)\right]}$$

In this equation, the term $\mathrm{GFR} \times f_{\mathrm{ub}}$ represents the renal excretion of unbound parent compound in blood by glomerular filtration, where GFR is the glomerular filtration rate, which is about 6.7 l/h in human adults, and f_{ub} is the fraction of the drug in the blood that is unbound (free). The second term in the denominator is hepatic clearance, where Q_{L} is liver blood flow (typically on the order of 90 l/h in adults) and $\mathrm{Cl}_{\mathrm{int}}$ is the intrinsic metabolic clearance for first-order conditions of metabolism in the liver at low concentrations.

PBPK Modeling

The basic approach to PBPK model development is illustrated in Figure 3.5. The process of model development begins with the identification of the chemical exposure and toxic effect of concern, as well as the species and target tissue in which the toxic effect is observed. Literature evaluation involves the integration of available information about (i) the mechanism of toxicity, (ii) the pathways of chemical metabolism, (iii) the nature of the toxic chemical species (i.e., whether the parent chemical, a stable metabolite, or a reactive intermediate produced during metabolism is responsible for the toxicity), (iv) the processes involved in absorption, transport, and excretion, (v) the tissue partitioning and binding characteristics of the chemical and its metabolites, and (vi) the physiological parameters (i.e., tissue weights and blood flow rates) for the species of concern (i.e., the experimental species and the human). Using this information, the investigator develops a PBPK model that expresses mathematically a conception of the animal/chemical system. In the spirit of the scientific method, model building is an iterative process. The literature includes many examples of successful PBPK models for a wide variety of compounds that provide a wealth of insight into various aspects of the PBPK modeling process. These published models should be consulted for further detail on the approach for applying the PBPK methodology in specific cases. This chapter will discuss general considerations in model development, validation, and application.

In the model, the various time-dependent chemical transport and metabolic processes are described as a system of simultaneous differential equations. Generally, these differential equations describe the rate of change in the amount of chemical in a particular compartment (tissue, blood) over time. Integration of these differential equations gives an estimate of the amount of chemical in the tissue at any one time, which then allows the calculation of chemical concentration in the tissue using real tissue volumes. As an example, the differential equation defining the liver compartment in Figure 3.1 is shown here:

$$dA_L/dt = Q_L \times (C_{Art} - C_L/P_L) - (V_{\max} \times C_L/P_L)/(K_M + C_L/P_L)$$

where A_{L} = the amount of chemical in the liver (mg); C_{Art} = the concentration of chemical in the arterial blood (mg/l); C_{L} = the concentration of chemical in the liver (mg/l); Q_{L} = the total (arterial plus portal) blood flow to the liver (l/h); P_{L} = the liver:blood partition coefficient (concentration ratio at equilibrium); $V_{\max}$ = the maximum rate of metabolism (mg/h); K_{m} = the affinity (concentration at half-maximum rate of metabolism) (mg/l).

PBPK Model Structure The specific structure of a particular model is driven by the need to estimate the appropriate measure of tissue dose under the various exposure conditions of concern in both experimental animals and humans. Before the model can be used in risk assessment, it must be validated against kinetic, metabolic, and toxicity data and, in many cases, refined based on comparison with the experimental results. The model itself is frequently used to help design critical experiments to collect data needed for its own validation. Refinement of the model to incorporate additional insights gained from comparison with

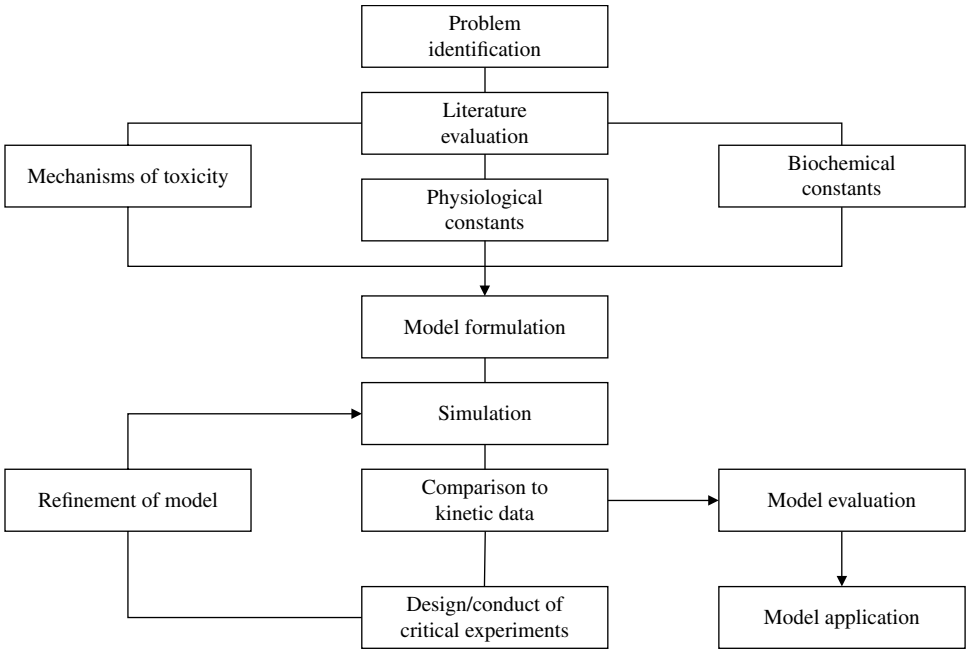


FIGURE 3.5 PBPK modeling approach.

experimental data yields a model that can be used for quantitative extrapolation well beyond the range of experimental conditions on which it was based.

There is no easy rule for determining the structure and level of complexity needed in a particular modeling application. For example, model elements such as inhalation and fat storage, which are important for a volatile, lipophilic chemical such as styrene, do not need to be considered in the case of a nonvolatile, water-soluble compound. Similarly, while kidney excretion and enterohepatic recirculation are important determinants of the kinetics for many compounds, they are not needed in the model of styrene. As another example, the simple description of inhalation uptake as a one-compartment gas exchange (Figure 3.1) may be adequate for some model applications, as in the case of modeling the systemic uptake of a lipophilic vapor like styrene. However, a more complicated description is required in the case of water-soluble vapors, to account for a "wash-in, wash-out" effect (chemical absorbed in the upper respiratory tract during inhalation can be reentrained in exhaled air during the second half of the breathing cycle). Thus, the decision of which elements to include in the model structure for a specific chemical and application entails striking a balance between two primary criteria: (1) parsimony (keeping it as simple as possible) and (2) biological plausibility (adequately describing the important determinants of the chemical kinetics).

After deciding which compartments to include in the model, decisions must be made regarding the description of the chemical kinetics, that is, transfer between blood and tissues, metabolism, and so on. Chemical transfer between the blood and tissue compartments may be governed by passive diffusion (flow- or diffusion-limited) or active transport. Many published PBPK models are flow-limited; that is, they assume that the rate of tissue uptake of the chemical is limited only by the flow of the chemical to the tissue in the blood. While this assumption is generally reasonable, for some chemicals and tissues the uptake may instead be limited by other factors such as diffusion. Examples of tissues for which diffusion-limited transport has often been described include the skin, placenta, mammary glands, brain, and fat. If there is evidence that the movement of a chemical between the blood and a tissue is limited by diffusion, a two-compartment description of the tissue can be used with a "shallow" exchange compartment in communication with the blood and a diffusion-limited "deep" compartment. Some chemicals may be transported against the concentration gradient through energy-dependent processes, rather than diffusion. These processes can result in high tissue:blood ratios for the chemical concentration and are sometimes limited by the availability of transporter proteins leading to saturation of chemical transport at high doses. Such saturable processes are often well-described using Michealis–Menten–type kinetics as discussed earlier for metabolism.

The liver is frequently the primary site of metabolism, though other tissues such as the kidney, placenta, lung, skin, and blood may also be important metabolism sites depending on the chemical. Metabolism may be described as occurring through a linear (first-order) pathway using a rate constant (k_F: h^{-1}) or a saturable (Michealis–Menten) pathway with capacity V_{max} (mg/h) and affinity K_m (mg/l). If necessary, the

PK of the resulting metabolite may also be explicitly described in the model. Metabolite kinetics are typically described when the metabolite is the chemical species causing the observed toxicity. The same considerations that drive decisions regarding the level of complexity of the PBPK model for the parent chemical must also be applied for each metabolite model. As in the case of the parent chemical, the most important consideration is the purpose of the model. If the concern is direct parent chemical toxicity and the chemical is detoxified by metabolism, then there may be no need for a description of metabolism beyond its role in parent chemical clearance. If reactive intermediates produced during the metabolism are responsible for observed toxicity, a very simple description of the metabolic pathways might be adequate. On the other hand, if one or more of the metabolites are considered to be responsible for the toxicity of a chemical, it may be necessary to provide a more complete description of the kinetics of the metabolites themselves.

Other processes that may have significant impact on the chemical kinetics, and may be included in the model, are protein binding and excretion. Protein binding in the blood reduces the amount of free chemical available for distribution into the tissues or clearance via excretion. Binding within tissues may lead to dose- and time-dependent accumulation, and may need to be described as a saturable process. Clearance may occur through urinary or fecal excretion, exhaled air, or even through loss via hair (as in the case of mercury). This loss may often be successfully described using first-order clearance terms. However, more elaborate descriptions are sometimes required for chemicals that are substrates for transport proteins that transfer the chemical against a concentration gradient. Some transporters in the kidney and bile can increase the clearance of xenobiotics, while others, such as those responsible for reabsorption, may decrease clearance.

PBPK Model Parameters Estimates of the various physiological parameters needed in PBPK models are available from a number of sources in the literature, particularly for the human, monkey, dog, rat, and mouse. Table 3.1 shows typical values of a number of physiological parameters in adult animals.

Estimates for the same physiological parameter often vary widely across sources, due both to experimental differences and to differences in the animals examined (age, strain, activity). Ventilation rates and blood flow rates are particularly sensitive to the level of activity. Data on some important tissues are relatively limited, particularly in the case of fat tissues.

Many biochemical parameters may be measured directly from *in vitro* studies. For volatile chemicals, for example, partition coefficients may be measured using a relatively simple *in vitro* technique known as vial equilibration. Partition coefficients for nonvolatile compounds are not as easily measured *in vitro*; however, and are therefore often estimated by comparing tissue:blood levels from *in vivo* studies. Metabolism parameters can be obtained from parent chemical disappearance (or metabolite formation) curves in cell suspensions, tissue homogenates, or microsomal fractions. Determination of urinary metabolites after *in vivo* exposure can also be useful for estimating metabolism parameters in some cases.

In many cases, important parameter values needed for a PBPK model may not be available in the literature. In such cases it is necessary to measure them in new experiments, to estimate them by QSAR techniques, or to identify them by optimizing the fit of the model to an informative data set. The process of adjusting a subset of the model's kinetic parameters to achieve the best agreement of the model simulation with measured tissue chemical concentrations is called "fitting" the model. An example of a case where fitting the model to kinetic data is the only practical approach for parameter estimation is the attempt to describe enterohepatic recirculation; that is, when a compound or its metabolite is transferred into the bile and subsequently reabsorbed from the intestine. Because enterohepatic recirculation is the result of several kinetic processes in the liver, bile, and GI tract, there is no system for measuring the process *in vitro*. As a result, model parameters for enterohepatic recirculation are generally determined by fitting the model to *in vivo* kinetic data. Even in cases where initial estimates of a particular parameter value can be obtained from other sources, it may be still desirable to refine the estimate by fitting *in vivo* data with the model. Of course, being able to uniquely identify parameters from a kinetic data set rests on two key assumptions: (1) the kinetic behavior of the compound under the conditions in which the data was collected is informative regarding the parameters being estimated, and (2) other parameters in the model that could influence the observed kinetics have been determined by other means and are held fixed or otherwise constrained during the estimation process.

Target Tissue Considerations Typically, a PBPK model used in risk assessment applications will include compartments for any tissues in which toxicity has been observed with a given compound (i.e., target tissues). The description of the target tissue may in some cases need to be fairly complicated, including features such as *in situ* metabolism, binding, and PD processes (e.g., upregulation of metabolizing enzymes) in order to provide a realistic measure of biologically effective tissue exposure.

A fundamental issue in determining the nature of the target tissue description required is identifying the active form of the compound. A compound may produce an effect directly through its interaction with tissue constituents or indirectly through a metabolite. Liver toxicity, in particular, is often caused by metabolism of the parent compound to reactive (short-lived) metabolites. Circulating metabolites may also lead to adverse effects in nonmetabolizing tissues.

TABLE 3.1 "Typical" Physiological Parameters for PBPK Models

Species	Mouse	Rat	Monkey	Human
Ventilation				
Alveolar (l/h–1 kg)[a]	29.0[b]	15.0[b]	15.0[b]	15.0[b]
Blood flows				
Total (l/h–1 kg)[a]	16.5[c]	15.0[c]	15.0[c]	15.0[c]
Muscle (fraction)	0.18	0.18	0.18	0.18
Skin (fraction)	0.07	0.08	0.06	0.06
Fat (fraction)	0.03	0.06	0.05	0.05
Liver (arterial) (fraction)	0.035	0.03	0.065	0.07
Gut (portal) (fraction)	0.165	0.18	0.185	0.19
Other organs (fraction)	0.52	0.47	0.46	0.45
Tissue volumes				
Body weight (kg)	0.02	0.3	4.0	80.0
Body water (fraction)	0.65	0.65	0.65	0.65
Plasma (fraction)	0.04	0.04	0.04	0.04
RBCs (fraction)	0.03	0.03	0.03	0.03
Muscle (fraction)	0.34	0.36	0.48	0.33
Skin (fraction)	0.17	0.195	0.11	0.11
Fat (fraction)	0.10[d]	0.07[d]	0.05[d]	0.21
Liver (fraction)	0.046	0.037	0.027	0.023
Gut tissue (fraction)	0.031	0.033	0.045	0.045
Other organs (fraction)	0.049	0.031	0.039	0.039
Intestinal lumen (fraction)	0.054	0.058	0.053	0.053

[a] Scaled allometrically: $QC = QCC \times BW^{75}$.
[b] Varies significantly with activity level (range: 15–40).
[c] Varies with activity level (range: 15–25).
[d] Varies substantially (lower in young animals, higher in older animals).

The specific nature of the relationship between tissue exposure and response depends on the mechanism, or mode of action, involved. Rapidly reversible effects may result primarily from acute compound concentrations in the tissue, while longer-term effects may depend on both the concentration and duration of the exposure. In fact, the appropriate measure of tissue exposure for one toxic effect of a compound may even be different from the appropriate measure for another of its effects. For example, the mitogenic effect of a compound may depend on the prolonged maintenance of a relatively high concentration sufficient to occupy a receptor in the target tissue, while cytoxicity may result from transient, high rates of metabolism occurring shortly after dosing. In such a case, PBPK modeling of the concentration time course in the target tissue for different dosing routes or regimens might be necessary. For developmental toxicity, windows of susceptibility must also be considered. That is, the fetus may be more susceptible to chemical toxicity during specific times in gestation that are associated with important developmental events (e.g., implantation, neural tube closure). Thus, the evaluation of the various modes of action for the beneficial and toxic effects of a compound is the most important step in a PK analysis and a principal determinant of the structure and level of detail that will be required in the PBPK model.

Interpreting a PBPK Model Diagram As described in the previous sections, the process of developing a PBPK model begins by determining the essential structure of the model based on the information available on the compound's toxicity, mechanism of action, and PK properties. The results of this step can usually be summarized by an initial model diagram, such as that depicted in Figure 3.1. In fact, a well-constructed model diagram, together with a table of the input parameter values and their definitions, is all that an accomplished modeler should need in order to recreate the mathematical equations defining a PBPK model. In general, there should be a one-to-one correspondence of the boxes in the diagram to the mass balance equations (or steady-state approximations) in the model. Similarly, the arrows in the diagram correspond to the transport or metabolism processes in the model. Each of the arrows connecting the boxes in the diagram should correspond to one of the terms in the mass balance equations for both of the compartments it connects, with the direction of the arrow pointing from the compartment in which the term is negative to the compartment in which it is positive. Arrows only connected to a single compartment, which represent uptake and excretion processes, are interpreted similarly.

Running a PBPK Model The previous sections have focused on the process of designing the PBPK model structure needed for a particular application. At this point

the model consists of a number of mathematical equations—differential equations describing the mass balance for each of the compartments and algebraic equations describing other relationships between model variables. The next step in model development is the coding of the mathematical form of the model into a form which can be executed on a computer. There are many options available for performing this process, ranging from programming languages such as Fortran, C, and MatLab to more user-friendly simulation software packages such as AcslX and Berkeley Madonna.

Evaluating a PBPK Model Model evaluation considers the ability of the model to predict the chemical's kinetic behavior under conditions that test the principal aspects of the underlying model structure. This is generally performed by running the model and attempting to predict measured tissue concentrations from data sets that were not used for model development. Goodness of fit, or correspondence between the model-predicted concentrations and measured concentrations, can be evaluated visually (subjective evaluation) or quantitatively through the use of automated algorithms that are available in most simulation software packages. While quantitative tests of goodness of fit may often be a useful aspect of the evaluation process, the more important consideration may be the ability of the model to provide an accurate prediction of the general behavior of the data in the intended application. Thus, if the model shows some deviation from measured concentrations, yet can consistently reproduce the trend of the data (biphasic clearance, saturation of metabolism, etc.) there will be greater confidence in the suitability of the model structure than a model that fits a portion of the data flawlessly. Indeed, the demand that the PBPK model fit a variety of data with a consistent set of parameters limits its ability to provide an optimal fit to a specific set of experimental data. For example, a PBPK model of a compound with saturable metabolism is required to reproduce both the high- and low-concentration behaviors, which appear qualitatively different, using the same parameter values. If one were independently fitting single curves with a model, different parameter values might provide better fits at each concentration, but would be relatively uninformative for extrapolation.

Ideally, model performance should be evaluated against data in the species, tissues and exposure scenarios of concern to risk assessors. However, it is not always possible to collect the data needed for such evaluation, particularly in humans. Where only some aspects of the model can be evaluated, it is particularly important to assess the uncertainty associated with those aspects that are untested. For example, a model of a chemical and its metabolites that is intended for use in cross-species extrapolation to humans would preferably be verified using data in different species, including humans, for both the parent chemical and the metabolites. If only parent chemical data is available in humans, the correspondence of metabolite predictions with data in several animal species could be used as a surrogate, but this deficiency should be carefully considered when applying the model to predict human metabolism. One of the values of biologically based modeling is the ability to use *in vitro* data to set model parameters, such as enzyme activity and substrate binding assays, which would improve the quantitative prediction of toxicity in humans from animal experiments.

Finally, it is important to remember that in addition to comparing model predictions to experimental data, model evaluation involves assessing the plausibility of the model structure and parameters and the confidence that can be placed in extrapolations performed by the model. This aspect of model evaluation is particularly important in the case of applications in risk assessment, where it is necessary to assess the uncertainty associated with risk estimates calculated with the model.

3.3 APPLICATIONS OF TOXICOKINETICS

Risk Assessment

The process of assessing health risks associated with human exposure to environmental chemicals inevitably relies on a number of assumptions, estimates, and rationalizations. Some of the greatest challenges in risk assessment result from the need to extrapolate from the conditions in the studies providing evidence of the toxicity of the chemical to the anticipated conditions of human exposure in the environment or workplace. For risk assessments based on animal data, the most obvious extrapolation that must be performed is from the tested animal species to humans. However, other extrapolations are also often required: from high dose to low dose, from one exposure route to another, and from one exposure time frame to another. PBPK modeling provides a powerful method for increasing the accuracy of these extrapolations. The inherent capabilities of PBPK modeling are particularly advantageous for cross-species extrapolation: physiological and biochemical parameters in the model can be changed from those for the test species to those which are appropriate for humans in order to provide a biologically meaningful animal to human extrapolation. Nonetheless, a full PBPK model may not always be necessary to support a PK risk assessment; in some cases (particularly those where human data are available) only a simple compartmental PK description is needed.

Simple PK approaches have sometimes been used by regulatory agencies in cancer risk assessment. However, the first case where an agency used a full PBPK approach was in the U.S. Environmental Protection Agency's (USEPA) revision of its inhalation cancer risk assessment for methylene

chloride. In 1989, the USEPA revised the inhalation unit risk and risk-specific air concentrations for methylene chloride in its Integrated Risk Information System (IRIS) database, citing a published PBPK model for the chemical. The resulting risk estimates were lower than those obtained by the USEPA's default approach by more than a factor of 10. That is, using the PBPK model, they determined that methylene chloride presented a substantially lower risk to humans than was predicted by the standard risk assessment calculations. Subsequently, an adaptation of the same PBPK model was used by the Occupational Safety and Health Administration in their rulemaking for a permissible exposure level (PEL) for methylene chloride. PBPK modeling has since become standard practice in risk assessment.

The ultimate aim of using PK modeling in risk assessment is to provide a measure of dose (dose metric) that better represents the "biologically effective dose"; that is, the dose that causally relates to the toxic outcome. The improved dose metric can then be used in place of traditional dose metrics (such as inhaled air concentrations or absorbed dose) in an appropriate dose–response model to provide a more accurate extrapolation to the human exposure conditions of concern. Implicit in any application of PK to risk assessment is the assumption that the toxic effects in the target tissue must be related to the concentration of the active form(s) of the chemical in that tissue. Moreover, in the absence of PD differences between animal species, it is expected that similar responses will be produced at equivalent tissue exposures regardless of species, exposure route, or experimental regimen. The motivation for applying PK in risk assessment, then, is the expectation that the observed effects of a chemical will be more simply and directly related to a measure of target tissue exposure than to a measure of administered dose.

The specific nature of the relationship between target tissue exposure and response depends on the chemical mechanism of toxicity, or mode of action, involved. Many short-term, rapidly reversible toxic effects, such as acute skin irritation or acute neurological effects, may result primarily from the current concentration of the chemical in the tissue. In such cases, the likelihood of toxicity from a particular exposure scenario can be conservatively estimated by the maximum concentration (C_{max}) achieved in the target tissue. On the other hand, the acute toxicity of highly reactive chemicals, as well as many longer-term toxic effects such as tissue necrosis and cancer, may be cumulative in nature, depending on both the concentration and duration of the exposure. A simple metric for such cases is the AUC in the tissue, which is defined mathematically as the integral of the concentration over time. This mathematical form implicitly assumes that the effect of the chemical on the tissue is linear over both concentration and time. The use of the AUC represents an extension of "Haber's Law," a concept developed from observations of the effects of chemical warfare gases that toxicity is proportional to the product of the concentration and time of exposure ($C \times T$). As stated previously, however, fetal effects will be determined by the chemical concentration and duration of exposure, as well as the gestational time period in which the exposure occurs.

Another complicating factor in describing the relationship between chemical concentration and tissue response is the need to determine the toxicologically active form of the chemical. In some cases, a chemical may produce a toxic effect directly, either through its reaction with tissue constituents (e.g., ethylene oxide) or by its binding to cellular control elements (e.g., dioxin). Often, however, it is the metabolism of the chemical that leads to its toxicity. In this case, toxicity may result primarily from reactive intermediates produced during the process of metabolism (e.g., chlorovinyl epoxide produced from the metabolism of vinyl chloride (VC)) or from the toxic effects of stable metabolites (e.g., trichloroacetic acid produced from the metabolism of trichloroethylene). The selection of the dose metric, that is, the active chemical form for which tissue exposure should be determined and the nature of the measure to be used, for example, peak concentration (C_{max}) or AUC, is the most important step in applying PK in risk assessment.

Whether intended or not, any dose metric will be consistent with the modes of action for some chemicals, and not for others. The USEPA, in a joint effort with scientists from several other agencies, prepared a review paper on cross-species extrapolation in cancer risk assessment, which concluded that "tissues experiencing equal average concentrations of the carcinogenic moiety over a full lifetime should be presumed to have equal lifetime cancer risk." The use of the term "carcinogenic moiety" in this statement reflects the concern that the dose metric should be representative of the active form of the chemical. For example, the use of the lifetime average daily concentration for the parent chemical might be appropriate for a directly genotoxic chemical such as ethylene oxide, which is detoxified by metabolism; however, it would not be appropriate for a chemical like VC, which requires metabolic activation to be genotoxic. In the latter case, increasing metabolism would increase the exposure to the genotoxic species but would decrease a dose metric based on parent chemical concentration. In such a case, where a reactive species produced during the metabolism of a chemical is responsible for its carcinogenicity, an appropriate cancer dose metric would be the lifetime average daily amount of metabolism in the target tissue divided by the volume of the tissue, as was used in the PK risk assessment for methylene chloride. Similar considerations apply in the case of noncancer risk assessment, except that the dose metrics are only averaged over the duration of the exposure (acute, subchronic, or chronic), not over a full lifetime.

The plausibility of a given dose metric is determined primarily by two factors: (1) its consistency with available information on the mode of action (mechanism of toxicity),

and (2) the consistency of its dose–response with that of the end point of concern. The first factor has been discussed earlier; the second factor refers both to evaluating the dose metric's ability to linearize the dose–response for the associated end point within a study (internal consistency), and to its ability to demonstrate a consistent quantitative relationship of dose metrics for positive versus negative exposures, regardless of differences in exposure scenario, route, and species (external consistency).

Example of Cancer Risk Assessment with a Toxicokinetic Model: Vinyl Chloride When it became evident that VC was carcinogenic both in animals and in humans, many of its uses were discontinued. The current use of VC is limited to serving as a chemical precursor in the production of such materials as polyvinyl chloride (PVC) and copolymer resins. However, VC is also produced from the biodegradation of another environmental chemical—trichloroethylene—by bacteria in the soil. Thus, past spills of trichloroethylene may lead to current or future exposures of the public to VC in drinking water or other environmental sources. The previous potency estimates for VC published by the USEPA did not quantitatively incorporate PK information on VC into the risk calculations. To provide a more accurate assessment of human risk from exposure to VC, a PBPK model was developed that describes the uptake, distribution, and metabolism of VC in mice, rats, hamsters, and humans following inhalation or oral exposure. The PBPK model was used to predict the total production of reactive metabolites from VC both in the animal bioassays and in human exposure scenarios. These measures of internal exposure were then used to predict the risk associated with lifetime exposure to VC in air or drinking water.

The PK risk assessment for VC demonstrated all of the attributes of an effective dose metric. First, the form of the metric (total daily metabolism divided by the volume of the liver) was consistent with the mode of action for the end point of concern (liver tumors), which involves DNA adduct formation by a highly reactive chloroethylene epoxide produced from the metabolism of VC. Second, while the dose–response for liver tumors versus exposure concentration of VC is highly nonlinear, with a plateau at several hundred ppm, the dose–response for liver tumors versus the metabolized dose metric is essentially linear from 1 to 6000 ppm. Finally, and most impressively, when the potency of VC liver carcinogenicity was expressed in terms of the metabolized dose metric, essentially the same potency was calculated from both inhalation and oral studies in the mouse and rat, as well as from occupational inhalation exposures in the human. Thus, PBPK models can support more accurate prediction of air and water concentrations that are likely to significantly increase the risk of cancer in the human population based on animal studies by eliminating factors (species differences in metabolism) that could confound such extrapolations.

Chemical-Specific Adjustment Factors

Risk assessments have typically applied default factors to account for uncertainty regarding animal to human extrapolation and human variability. That is, when significant uncertainty exists in an aspect of the risk estimate, factors are applied in order to make it more conservative. These uncertainty factors are not based on data, but rather on the assumption that decreasing exposure by a factor of 10–1000 will account for the different types of uncertainty (interspecies or intraspecies differences in TK or TD) in the risk assessment process. Significant progress has been made in recent years in refining this approach beyond the use of default uncertainty factors. An important step forward in the development of approaches for incorporating chemical-specific data in risk assessment is the recent guidance from the International Programme for Chemical Safety (IPCS) addressing the data requirements for replacing default uncertainty factors with chemical-specific adjustment factors (CSAFs). The IPCS CSAF approach breaks the inter- and intraspecies uncertainty factors into TK and TD components, each of which can be replaced by a CSAF if adequate chemical-specific data are available. The TK factor for interspecies differences (AK_{UF}) represents the ratio of the external exposures in humans and animals that would produce the identical internal (target tissue) exposures. Similarly, the TK factor for human variability (HK_{UF}) represents the ratio of the doses in average and sensitive individuals that would produce the identical internal (target tissue) exposure. Depending on the data available for the chemical, the magnitude of the adjustment factor for TK may be calculated based on a variety of biokinetic factors, such as the clearance of the chemical or the AUC for the chemical. For example, a cross-species TK adjustment factor for boric acid has been estimated by USEPA on the basis of the ratio of glomerular filtration rates in animals and humans. PBPK models can also be used to estimate the adjustment factors for TK, as described in the example here.

Example of the Calculation of a CSAF Using a PBPK Model The calculation of a CSAF for interspecies differences in TK, AK_{UF}, for 2-butoxyethanol provides a good example of the approach and considerations required for the IPCS methodology. In the case of 2-butoxyethanol, several PBPK models had been developed that could be used to determine the cross-species adjustment for TK. In fact, it would be difficult to determine the AK_{UF} in this case without a PBPK model. This is because the best animal data available to support the calculation of an AK_{UF} consist of AUCs of 2-butoxyacetic acid in the blood of rats exposed to 2-butoxyethanol by inhalation for 6 h; however, the AUCs were reported for the postexposure period only. Therefore, it was necessary to estimate the total AUC using a PBPK model, by integrating the predicted concentration of 2-butoxyacetic acid in venous blood both

during and following an inhalation exposure of 6 h. Using the rodent PBPK model, Health Canada determined that the AUC during the exposure period was actually on the same order as the AUC reported for the postexposure period. Thus, use of the reported AUCs would result in a factor of two errors in estimating the AK_{UF}. Use of the human data in the calculation of an AK_{UF} was also problematic, because the exposures were conducted under exercising conditions. Analyses performed with the human PBPK model indicated that the uptake of the parent compound is linearly related to the ventilation rate. Therefore, the AUC value in this study was adjusted to account for working versus resting conditions using the results of the human PBPK model. Note that the effect of ventilation rate on the highly soluble 2-butoxyethanol contrasts with the case of poorly soluble, lipophilic compounds, where ventilation rate has little impact on uptake. Thus, different adjustment factors would be calculated for each of these compounds, rather than using the same default adjustment factor in both cases.

Interindividual Variability

One of the more challenging issues that must be considered in performing a human health risk assessment is the heterogeneity among humans. This heterogeneity is produced by interindividual variations in physiology, biochemistry, and molecular biology, reflecting both genetic and environmental factors, and results in differences among individuals in the biologically effective tissue dose associated with a given environmental exposure (PK) as well as in the response to a given tissue dose (PD). This interindividual variability is not as evident in animal studies, where breeding and environmental controls limit within-study differences. Because the parameters in a PBPK model have a direct biological correspondence, they provide a useful framework for determining the impact of observed variations in physiological and biochemical factors on the population variability in dosimetry within the context of a risk assessment for a particular chemical.

It is useful to consider the total variability among humans in terms of three contributing sources: (1) the variation across a population of "normal" individuals at the same age, for example, young adults; (2) the variation across the population resulting from their different ages, for example, infants or the elderly; and (3) the variation resulting from the existence of subpopulations that differ in some way from the "normal" population, for example, due to genetic polymorphisms. A fourth source of variability, health status, should also be considered, although it is frequently disregarded in environmental risk assessment. To the extent that the variation in physiological and biochemical parameters across these population dimensions can be elucidated, PBPK models can be used together with Monte Carlo methods to integrate their effects on the *in vivo* kinetics of a chemical exposure and predict the resulting impact on the distribution of risks (as represented by target tissue doses) across the population.

Determinants of Impact There has sometimes been a tendency in risk assessments to use information on the variability of a specific parameter, such as inhalation rate or the *in vitro* activity of a particular enzyme, as the basis for expectations regarding the variability in dosimetry for *in vivo* exposures. However, whether or not the variation in a particular physiological or biochemical parameter will have a significant impact on *in vivo* dosimetry is a complex function of interacting factors. In particular, the structures of physiological and biochemical systems frequently involve parallel processes (e.g., blood flows, metabolic pathways, excretion processes), leading to compensation for the variation in a single factor. Moreover, physiological constraints may limit the *in vivo* impact of variability observed *in vitro*. For instance, high-affinity intrinsic clearance can result in essentially complete metabolism of all the chemical reaching the liver in the blood; under these conditions, variability in amount metabolized *in vivo* would be more a function of variability in liver blood flow than variability in metabolism *in vitro*. Thus it is often true that the whole (the *in vivo* variability in dosimetry) is less than the sum of its parts (the variability in each of the PK factors).

The dosimetric impact of variations in physiological factors also depends on the nature of the chemical causing the toxicity, including such physicochemical properties as reactivity, lipophilicity, water solubility, and volatility. For example, variations in inhalation rate will tend to have more impact on the uptake of a water soluble chemical such as isopropanol than on a relatively water insoluble chemical such as VC.

In addition, the impact of a particular factor on dosimetry also depends on the mode of action of the chemical; that is, how the chemical causes the effect of concern. Of particular importance is whether the toxicity results from exposure to the chemical itself, one of its stable, circulating metabolites, or a reactive intermediate produced during its metabolism.

Another key issue is whether the toxicity results from direct reaction with tissue constituents, from binding to a receptor, or from physical (e.g., solvent) effects on the tissue. To illustrate these considerations, one can contrast the acute neurotoxicity of many solvents (a physical effect of the chemicals themselves) with their chronic hepatotoxicity (produced by products of their metabolism). The most important PK factor in the acute toxicity of volatile solvents is the blood:air partition coefficient, and increasing metabolic clearance typically decreases toxicity. In contrast, the most important PK factors in the chronic toxicity are liver blood flow and metabolism, and increasing metabolic clearance typically increases toxicity.

Example: Age-Dependent Variability The following example illustrates the use of PBPK modeling to investigate the impact of PK variability on risk for the case of age-dependent PK. Specifically, the question being evaluated in this example is how normal changes in PK parameters from birth, through childhood and adulthood affect the dosimetry for environmental chemical exposures. To this end, a PBPK model was developed to simulate the physiological and biochemical changes in humans associated with growth and aging. In the age-dependent model, all physiological and biochemical parameters change over time based on data from the literature.

Figure 3.6 shows the results of using this age-dependent model to simulate continuous inhalation of isopropanol at air concentrations of 1 ppb, beginning at birth and continuing for 75 years. The model predicts that, for the same inhaled concentration, the blood concentrations of isopropanol and its metabolite, acetone, achieved during early life are significantly higher than those achieved during adulthood. The capacity for metabolism is reduced in the infant compared to older children and adults.

Quite a different behavior is predicted, however, for daily ingestion of perchloroethylene in drinking water. The exposure in this case is assumed to be a constant intake of perchloroethylene at a rate of 1 μg/kg/day throughout life. As shown in Figure 3.7, predicted concentrations of perchloroethylene and its major metabolite, trichloroacetic acid, are much lower during early life than during adulthood. In contrast to isopropanol, the most important determinant of the kinetics of perchloroethylene is its tendency to accumulate in fat, while changes in metabolism have little impact. It is also interesting to note that for this lipophilic chemical, different behaviors are predicted in males and females due to sex-dependent differences in body fat content.

Example: Genetic Polymorphism The next example demonstrates the use of PBPK modeling, together with Monte Carlo techniques, to evaluate the impact of a genetic polymorphism for metabolism. In this example, the polymorphism of interest is for the enzyme paraoxonase. The PBPK model used in the analysis describes exposure to parathion, its metabolism to paraoxon, and the inhibition of acetylcholinesterase by paraoxon. Paraoxonase is one of the enzymes responsible for the metabolic clearance of paraoxon. *In vitro* data on the two human alleles of paraoxonase (low and high activity) were used to develop distributions for the metabolism parameters in the PBPK model. Monte Carlo simulations were then performed to generate the resulting distribution of predicted blood concentrations of paraoxon across a population, considering the variability in other PK parameters. Figure 3.8 displays the predicted distribution for the time-integrated (area under the curve) blood concentrations of paraoxon (mg-h/l) across the total population (dark bars), as compared to the "normal" population (light bars, excluding individuals who are homozygous for the low activity allele), following exposure to parathion at a dose of 0.033 mg/kg. While the polymorphism does impact the distribution of blood concentrations, particularly at the higher internal exposures, the overall effect is relatively small when put in the perspective of the variability in other physiological and biochemical factors across the same population.

The overall PK variability across a population is a function of many chemical-specific, genetic, and physiological factors.

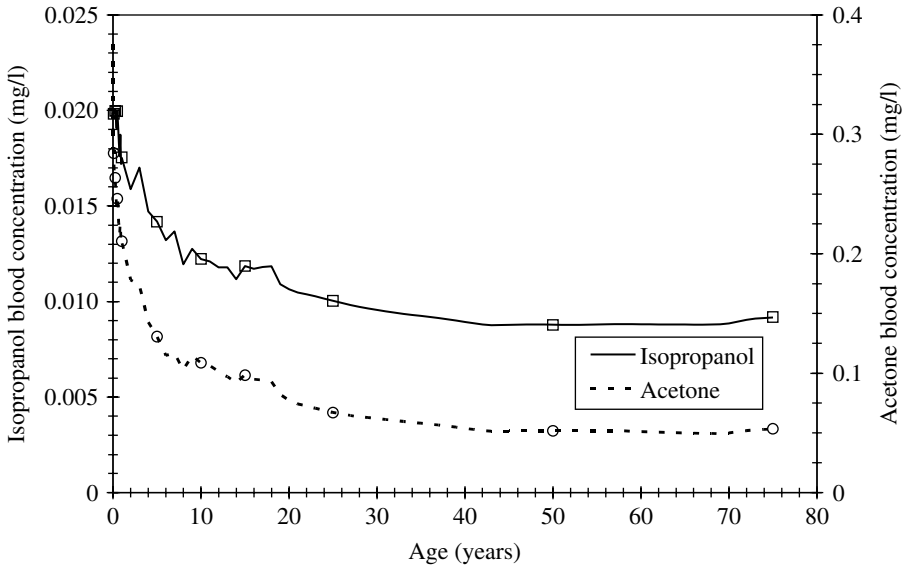


FIGURE 3.6 Blood concentrations of isopropanol (IPA) and its metabolite acetone as a function of age for continuous inhalation exposure at 1 ppb. *Source:* Adapted from Clewell et al. (2004).

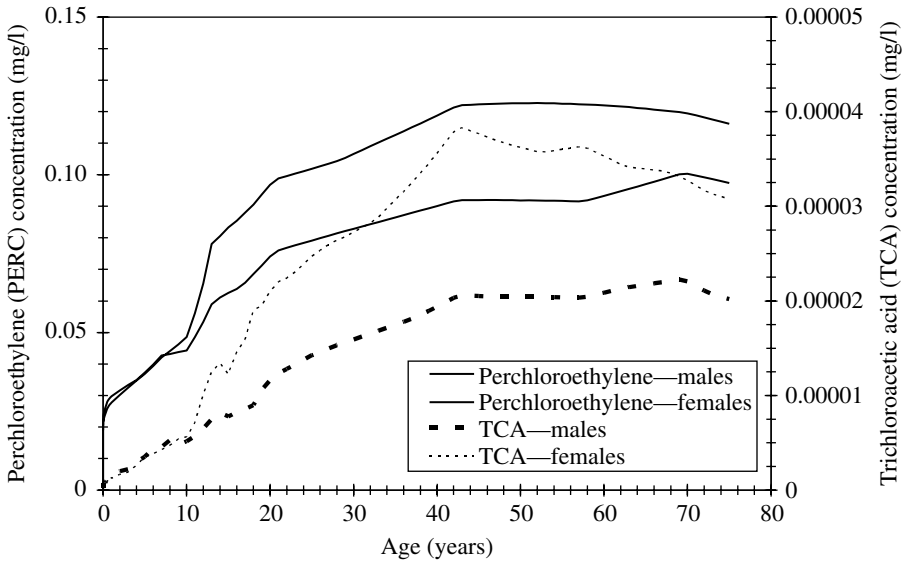


FIGURE 3.7 Blood concentrations of perchloroethylene (PERC) and its primary metabolite trichloroacetic acid (TCA) as a function of age for continuous oral exposure at 1 μg/kg/day. *Source:* Adapted from Clewell et al. (2004).

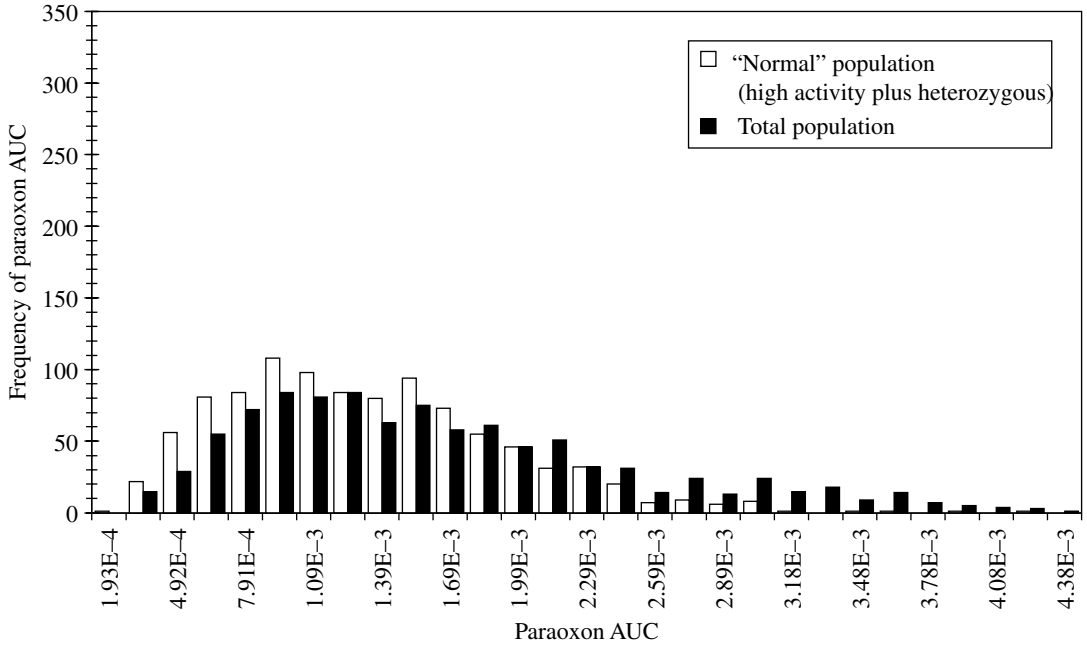


FIGURE 3.8 Paraoxon AUC distribution for the total population (Case 3, including low-activity allele) compared to the AUC distribution for the "normal" population (Case 2, using only high-activity and heterozygous alleles) for a parathion dose of 0.033 mg/kg. *Source:* Adapted from Gentry et al. (2002).

Due to the complex interactions among these factors, speculation regarding the extent of population variability on the basis of the observed variation in a single factor can be highly misleading. Analysis using PBPK modeling and Monte Carlo techniques provides a more reliable approach for estimating population PK variability. Analyses such as that described here for parathion can be used to develop quantitative CSAFs to replace default uncertainty factors for human PK variability. PBPK modeling can also be useful in a more qualitative sense, to determine whether there is reason for concern regarding a particular age group that might be more sensitive due to PK differences, as illustrated by the

example using the age-dependent model. Similar analyses can be performed to determine whether exposure during special life stages, such as gestation or lactation, represents a significant concern. PBPK modeling of this nature, coupled with parameter estimation using QSAR techniques and mechanistic information from genomic arrays, may prove particularly valuable in prioritizing testing requirements for new chemicals.

3.4 CASE STUDY: TOXICOKINETICS OF DIBUTYL PHTHALATE

Dibutyl phthalate (DBP) belongs to a family of dialkyl esters (phthalates) that are primarily used as softeners in plastic products, but may also be found in paint, glue, putty, pharmaceutical products, and cosmetics. Phthalates are not chemically bound to the plastics and, as a result, they may leach from the surface over time. Human exposure primarily occurs via the oral route, as phthalates from plastic containers may leach into the packaged foods or liquids, which are then ingested. In the case of DBP, dermal and inhalation exposures are negligible in comparison to oral absorption. Biomonitoring studies have identified the metabolites of DBP and several other phthalates in the urine of the U.S. population including adults, children, infants, and pregnant women. Estimated exposure to DBP is in the range of 0.001–0.004 mg/kg/day in the United States. The primary toxicological concern with respect to DBP exposure is the possibility of developmental effects following gestational exposure.

In hazard identification studies, high doses of DBP administered to pregnant rats (>100 mg/kg/day) interfere with sexual development of the male rat fetus. Specifically, DBP causes feminization of the male rat, which is characterized by reduced anogenital distance (AGD), nipple retention, hypospadias (malformed penis), delayed testes descent and other effects on the reproductive organs of male offspring. Many of these effects are associated with reduced testosterone production in the testes of the male fetus during gestation. The similarity between the steroidogenic pathway in rats (where toxicity has been observed) and humans and the widespread exposure to phthalates has led to a high degree of public concern about the potential for adverse effects in humans after perinatal exposure through medical devices and childcare products (i.e., bottles, pacifiers, and infant toys).

In order to estimate the risk associated with human exposures, an understanding of relative dose is required. Specifically, we must know (i) what concentration of the active chemical is required to cause effects in the target tissue, and (ii) how the human tissue concentrations compare to concentrations that are associated with negative effects in rats. In the case of DBP, the active metabolite (metabolite that directly interferes with testosterone-mediated development) is the monoester metabolite monobutyl phthalate (MBP), and the target tissue is the testis of the male fetus (where testosterone is produced). The primary goal of the TK studies and PBPK models described in the following case study was to determine concentration of MBP in the male fetal rat testes in exposure conditions where adverse effects on development have been observed and to better estimate the risk associated with human exposures.

Design of Kinetic Studies

When designing TK studies several important decisions must be made regarding (i) the species and strain of animals to be used, (ii) the dose regimen, (iii) the time points for tissue collection, (iv) the tissues that will be collected for chemical analysis, and (v) appropriate methods for tissue sample preparation and storage. The first four decisions will primarily be determined by the published data for the chemical of interest. It is important that the kinetic studies be performed using the animal model and dose regimen that have been associated with toxicity in hazard studies. If a toxic effect has been identified in rabbits, for example, a kinetic study in the mouse or rat would not be appropriate, as no information could be gained on the relationship of target tissue dose to toxic outcome. Likewise, if effects have been identified in rats after inhalation, little will be learned by studying chemical kinetics following i.v. administration. Dosing regimen and time points will be influenced by prior knowledge of the kinetics for the chemical (if available) or for other similar compounds. The estimated half-life is often used in determining whether a chemical should be tested after a single dose or after repeated doses: chemicals with long half-lives (>6 h) in the test species will take more than 1 day to come to steady state. Thus, it is often important to test in both acute and subchronic (several days to weeks) scenarios. The half-life is also useful in determining the time points to be tested. It is generally accepted that in order to obtain accurate estimates of kinetic parameters, including clearance and AUC, the kinetic time course should be followed out to five to six half-lives. Thus, a chemical with an estimated half-life of 4 h will generally be studied out to 24 h. Additional time points are collected to determine the concentrations around T_{max}.

Several factors are important in determining which tissues will be collected and analyzed. In general, blood, plasma or serum is always necessary. If a chemical is known to accumulate in blood cells, however, it may be important to collect whole blood and analyze the plasma and RBCs separately. Tissues of interest for TK studies include those responsible for chemical metabolism (liver, kidney, lungs), sites of storage (fat), routes of elimination (feces, urine), and target tissues.

In the case of DBP, several hazard and kinetic studies are available in the literature that could help in designing the optimal kinetic study. The design of the gestation studies is described here, together with the reasoning behind experimental design decisions.

Animal Model Reproductive effects of DBP and similar phthalates had been identified in rats of more than one strain (Sprague-Dawley, Wistar, Fisher), and effects in the fetus and adult offspring of exposed pregnant rats were generally consistent across the tested strains of rats. Studies in mice, however, were unable to recapitulate the effects observed in rats, indicating that the mouse is not the appropriate model for testing phthalates. Therefore, in order for the kinetic studies to provide useful information for the hazard studies, rats must also be used for the kinetic studies. As the majority of the effects studies had been performed in Sprague-Dawley rats, this same strain was used for TK studies.

Dose Regimen Despite the fact that the dietary route of exposure is more relevant for human exposure, the vast majority of developmental studies performed to evaluate the effect of DBP on the male reproductive tract use oral gavage administration. DBP is typically administered to the maternal rat daily during late gestation, starting at gestation day (GD) 12–14, and continuing through either the end of gestation (GD 21) or lactation. Days 12–21 of gestation have been clearly identified as the window of susceptibility for disruption of male rat sexual development by DBP and similar phthalates, that is, the fetal rat is more susceptible to DBP if the exposure occurs in late gestation (after day 12) rather than early gestation. GD 19 was further identified as the most sensitive time for measuring testosterone inhibition. In order to recapitulate the dose regimen associated with testosterone-mediated effects, pregnant rats were administered DBP via oral gavage from GD 12 to GD 19 and chemical kinetics were measured following the final dose. Dose levels (50, 100, and 500 mg/kg/day) were determined based on hazard studies, which had shown effects in the fetus at concentrations ranging from 50 to 500 mg/kg/day, and kinetic studies in adult rats that indicated differences in clearance kinetics between doses of 100 and 500 mg/kg/day. Since DBP is highly lipophilic; it was administered with corn oil as a vehicle.

Time Points for Tissue Collection Published studies in rats indicate that DBP is rapidly and completely hydrolyzed to the monoester metabolite (MBP) in the gut prior to absorption. MBP and its glucuronide conjugate (MBP-G) are the major metabolites in the blood, urine, and tissues. Furthermore, the MBP metabolite has a half-life of approximately 4 h in the rat. Based on this information, it is reasonable to follow the metabolite kinetics out to 24 h postdosing. C_{max} would be expected to occur prior to 4 h. To define the rate of uptake of MBP and identify C_{max}, tissue collection was performed at 0.25, 0.5, 0.75, 1, and 2 h postdosing. Later time points (4, 8, 12, and 24 h) were also included to characterize the clearance of the metabolites.

Tissues Collected for Analysis In addition to plasma, several tissues were collected for metabolite analysis. Maternal liver was collected, as it is the primary site of MBP metabolism (glucuronide conjugation). The primary purpose of the study, however, was to define fetal metabolite exposure. Maternal plasma, placenta, and fetal plasma were collected to characterize maternal–fetal transfer of the metabolites. Amniotic fluid was collected as it can play a role in determining fetal concentrations (amniotic fluid is ingested by the fetus). Fetal testes were also collected for metabolite analysis as they are the target tissue.

Methods for Tissue Storage and Analysis For some chemicals, it is necessary to take steps to prevent continued metabolism after removal of the tissue or collection of blood as many enzymes can retain their activity for a period of time after removal from the body. Methods for stopping enzymatic processes include adding chemical inhibitors of the enzymes to the samples or flash-freezing samples using liquid nitrogen. MBP is primarily metabolized in the liver through glucuronidation, though there is some oxidative metabolism by P450 enzymes. In order to minimize continued metabolism, the liver was placed in a vial and flash-frozen in liquid nitrogen immediately after removal.

In the case of the phthalates, special precautions were needed to ensure that samples were not contaminated with exogenous phthalates. Phthalates are widely used in plastics and many studies have reported the presence of MBP in samples in control animals (no DBP exposure). Due to the prevalence of the chemical in plastics, there are many points at which DBP may be inadvertently introduced to the experiment. To avoid this, phthalate-free plastics or silanized glass containers were used for sample storage. This is particularly important for the plasma, as phthalates have been shown to leach from plastic containers into stored blood and plasma even while frozen.

Analysis of Kinetic Studies

PK parameters for MBP and MBP-G in the maternal and fetal plasma were calculated using NCA with extravascular (not i.v.) administration. A comparison of kinetic parameters provided significant insight into the nonlinearity of the kinetics, that is, the differences in kinetic behavior (clearance, oral absorption, metabolism) observed at different doses. A linear response occurs when the rates of clearance, metabolism, and oral absorption do not change as dose increases. As such, we would expect the blood (or plasma) levels to increase linearly with administered dose, that is, a twofold increase in dose would yield a twofold change in

blood concentrations (AUC, C_{max}). Analysis of the kinetic parameters for DBP showed that this chemical does not have a linear relationship between administered dose and blood concentrations.

A simple method for testing whether the blood concentrations increase linearly with dose is to compare the dose-normalized AUC and C_{max} (i.e., the blood AUC or C_{max} divided by the administered dose; AUC/D, C_{max}/D). If the blood concentrations increase linearly with dose, then the AUC/D and C_{max}/D should remain constant across doses. In the case of DBP, AUC/D was approximately twofold higher in the plasma of the 100 and 500 mg/kg/day rats than those given 50 mg/kg/day, which is consistent with nonlinear kinetics. This nonlinear increase in blood concentrations was identified to be the result of decreased clearance of the monoester due to saturation of metabolism.

In addition to the nonlinearity in AUC/D, C_{max}/D, the time of peak concentration (T_{max}), and the mean residence time (MRT) also showed dose-dependent differences. C_{max}/D, T_{max}, and MRT were similar between the 50 and 100 groups, but differed in the 500 mg/kg/day group. The C_{max}/D was reduced nearly threefold at 500 mg/kg/day, while T_{max} and MRT were increased more than twofold in the 500 mg/kg/day group. The reproducible increase in T_{max} and corresponding decrease in C_{max}/D seen at the highest dose indicates that uptake of the monoester is dose-limited; that is, the chemical is absorbed less efficiently at higher doses.

The observed nonlinearity of DBP/MBP kinetics has important implications for the interpretation of hazard studies. For DBP, many of the more serious developmental effects have been seen only at doses at or above 500 mg/kg. Reduced AGD, for example, which is a source of concern due to its indication of impaired sexual development in both rats and humans, is only seen at doses greater than 250 mg/kg/day. This effect may in fact be a result of reduced clearance of MBP due to saturation of metabolism—a phenomenon that will not occur in humans as human exposure occurs at much lower levels than in rats. Therefore, it would be more appropriate to use effects seen at lower doses as markers of toxicity, because they occur within the range of linear kinetics.

Development of a PBPK Model

The combination of multiple nonlinear metabolic processes (discussed earlier) leads to rather complex kinetic behavior of DBP metabolites *in vivo*, making it difficult to predict a priori the tissue concentrations at a given dose. Inherent changes in physiology and biochemistry during pregnancy further compound this issue. Yet, in order to evaluate the risk associated with exposure to DBP, it is necessary to have a quantitative understanding of the dose–response of the active compound, MBP, at the target organ (e.g., fetal testis). Because PBPK models are based on physiological, biochemical, and mechanistic data, they allow us to estimate target site dosimetry while improving confidence in extrapolation between dose levels, exposure durations, and routes. Thus, we may more accurately predict tissue concentrations for exposure scenarios that have not been directly tested *in vivo*.

Model Structure Development of a model structure is an iterative process. The initial model structure is typically determined based on prior knowledge of the intended use (chemical mode of action and target tissues), important kinetic process driving chemical behavior (metabolism, clearance), and available data obtained from literature searches. Simulations are then performed and compared to kinetic data. Model parameters are adjusted to fit the experimental data (to the extent possible while following the constraints of realistic biology). As the model is tested against the data, it often becomes apparent that changes to the model structure are required to recapitulate the behavior of the measured data sets. The model structure is then refined, tested again against the data, and refined again if necessary. The model may also be used to design experiments that can help verify the proposed model structure and test mechanistic hypotheses.

The final structure of the DBP model (Figure 3.9) was the product of just such an iterative process. The initial structure was highly simplified compared to the final published model structure. The initial structure of the PBPK model for DBP was determined based on two primary considerations: its intended use and the available data. For DBP, the intended use was ultimately to aid risk assessment efforts by (i) accurately predicting fetal testes concentrations associated with acute and chronic exposures in the rat, (ii) providing a platform for extrapolating predictions to additional dosing scenarios (high dose to low dose, acute to chronic), and (iii) eventually predicting human tissue exposures. As such, it was necessary to include descriptions of both DBP and MBP, as exposures generally occur as DBP and MBP is the active metabolite. Likewise, it was necessary to include descriptions of both maternal and fetal kinetics, as exposure occurs via the maternal rat and the target tissue is the fetal testes.

The model structure is also heavily influenced by the availability of data. It is important to include all of the key processes controlling chemical kinetics and the tissue compartments that play pivotal roles in the distribution, elimination, or toxicity of a chemical. However, those who are designing the models must recognize that as the number of compartments in the PBPK model increases the number of input parameters also increases, and each of these parameters must be estimated from experimental data of some kind. Thus, the availability of data to support the model components will often provide practical limits to the modeler. It is often preferable to simplify models where data is scarce, in order to limit the number of unsupported parameters. As the majority of the more detailed kinetic data for DBP was collected in adult male rats, the initial model was

developed for the adult male where plentiful data existed for model parameterization and validation. The validated model was then extended to gestation as data became available.

Through rigorous testing of the model simulations, it was determined that significant refinement of the initial model structure would be necessary. For example, in order to track the chemical disposition across a wide range of doses, it was found that the downstream metabolites of MBP (MBP-G and oxidative metabolites of MBP; MBP-O) would need to be included in the model in addition to DBP and MBP. MBP-G was important to describing gut dynamics and the sustained presence of free MBP in the blood (a result of enterohepatic recirculation of MBP-G followed by conversion to free MBP in the GI). At higher doses, MBP disposition could be well described by accounting for DBP, MBP, and MBP-G disposition. At lower doses, however, a significant portion of the dose is subjected to oxidative metabolism at a high affinity compared to glucuronidation. Thus, in order to ensure that the model could predict both high and low doses, it was necessary to include the oxidative metabolic pathway. Additional refinements included building a more comprehensive description of chemical dynamics in the GI (enterohepatic recirculation), adding descriptions of MBP metabolism within the fetus, and so forth. The final model structure for DBP disposition in the pregnant and fetal rat is shown in Figure 3.9.

The final model (Figure 3.9) contains four interconnected submodels, each with the necessary amount of detail to adequately describe its chemical species: DBP, MBP, MBP-G, and the combined oxidative metabolites (MBP-O). The individual submodels interact at sites of metabolism (hydrolysis of the diester, glucuronidation, hydrolysis of the glucuronide, and oxidation). The models for each chemical species in the adult rat are described here, followed by the modifications made to describe gestation.

Adult Male Rat Model

Intact DBP

Enzymes responsible for the hydrolysis of DBP in the plasma and liver are described as first-order rates based on the experimental data. Hydrolysis of the DBP in the upper GI (stomach + small intestine), on the other hand, is described as a saturable process based on the *in vitro* and *in vivo* data indicating saturation of hydrolysis and oral uptake at the highest doses. Unmetabolized DBP is poorly absorbed in the gut or passed in to the lower intestine (G2) and cleared in the feces. Oral absorption and transport between the upper intestine (G1) and lower intestine (G2) compartments are described using first-order rates. Fecal excretion is described using a clearance rate (l/h). DBP that is taken up into the gut wall is passed to the liver via the portal blood where it is hydrolyzed, released into systemic circulation with the plasma, or excreted into the bile. Biliary transfer of DBP into the duodenum (enterohepatic recirculation) is modeled as a clearance rate from the liver to the upper intestine. Transport of DBP into the tissues is flow-limited.

Free MBP

Oral absorption is described as a first-order process and movement through the GI and fecal excretion are clearance rates. Unlike DBP, MBP is readily absorbed in the gut wall and passed to the liver via the portal blood. Glucuronidation and oxidation of free MBP in the liver are described using saturable kinetics (Michaelis–Menten). Free MBP may also be excreted into the bile or released into systemic circulation. Transfer of MBP in the bile is described in the same manner as DBP, using a clearance rate from the liver to the upper intestine (GC1). The MBP may then be reabsorbed or transported into the lower intestine. Absorption may occur in both the upper and lower intestine compartments. Free MBP that is not absorbed in the intestine is cleared via the feces. Transport of MBP into the tissues from the plasma is diffusion-limited.

MBP-glucuronide

MBP-G formed in the liver may be excreted into the bile or released into systemic circulation. Biliary transfer of MBP-G is modeled as described for free MBP earlier. MBP-G then travels through the intestine and is either hydrolyzed to MBP via β-glucuronidase in the lower intestine (GC2) or passed in the feces using first-order clearance rates. In the model, the hydrolysis of MBP-G is the rate limiting step for reabsorption of conjugated MBP from the bile. Distribution of MBP-G into the tissues is modeled using flow limitation, assuming distribution with body water. Urinary excretion of MBP-G is modeled as a first-order clearance rate from the plasma compartment.

Oxidative Metabolites of MBP

MBP-O formed by P450 metabolism in the liver is released into the body via the venous blood. A one-compartment volume of distribution model is used to describe the combined oxidative metabolites. Distribution is assumed to be the body water compartment. Urinary excretion is modeled as a first-order clearance rate from the central compartment. This model was kept relatively simple due to the lack of data for the oxidative metabolites of MBP.

Modifications for Gestation During gestation, both MBP and MBP-G are allowed to move freely between the arterial and placental plasma. Transfer of MBP between the placental plasma and the fetal plasma is diffusion-limited. Versions of the model that did or did not allow MBP-G to cross the placenta were tested (described later). Based on the fit of the model simulations to available data, the final model does not include placental transfer of MBP-G. Based on fetal MBP kinetic data, glucuronidation of MBP and hydrolysis of MBP-G are included in both the dam and the fetus. Transfer of MBP and MBP-G between the fetus and amniotic fluid are described as first-order processes. Transfer between the fetal plasma and testes tissue is diffusion-limited.

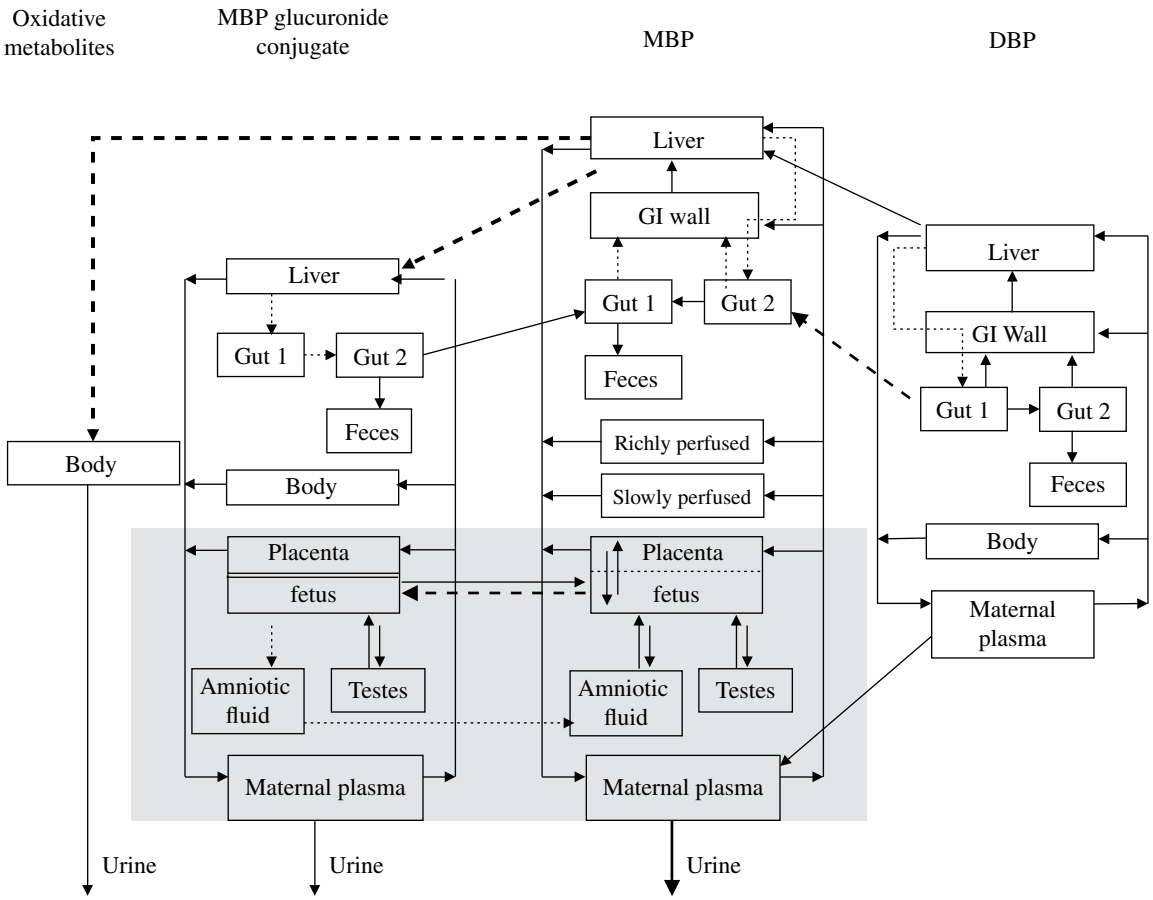


FIGURE 3.9 Model structure for DBP kinetics in the pregnant rat. *Dashed arrows* indicate first order processes and clearance rates. *Bold dashed arrows* represent saturable processes (i.e., metabolism). *Solid arrows* represent blood flows to the tissue compartments. The *solid arrows* represent flow-limited (DBP) and diffusion-limited (MBP, MBP-G) transport into the tissues. *Shaded area* illustrates description of maternal–fetal chemical transfer. *Source:* Adapted from Clewell et al. (2008).

Model Parameterization

Physiological Parameters Physiological parameters were obtained from measured values reported in the literature.

Allometric Scaling Kinetic parameters were scaled allometrically as is typical for intra- and interspecies extrapolation. PA, V_{max}, and clearance constants were scaled by $BW^{0.75}$, where BW is the body weight of the mother. With the exception of metabolism parameters, fetal parameters were scaled in a similar manner to the maternal parameters; PAs were scaled by $v_{fet}{}^{0.75}$, where v_{fet} is the volume of the fetus, and then multiplied by the total number of fetuses to obtain the value for the litter. Metabolism parameters (glucuronide conjugation and hydrolysis) were scaled by adjusting the adult value by the ratio of the fetal:adult liver weight, since the fetal liver weight is not linearly correlated with body weight. The value for the total litter was then calculated by multiplying by the number of fetuses.

Setting Kinetic Parameters Whenever possible, parameters were taken from published values or calculated from *in vitro* studies. Tissue:plasma partition coefficients were obtained from radiolabeled DBP studies and vial equilibration studies. The affinity constant (K_m) for hydrolysis in the gut was calculated from *in vitro* metabolism studies. Nonetheless, the lack of specific tissue and metabolism data required that many of the model parameters be fitted to *in vivo* kinetic data. For parameters that were not measured experimentally, values were estimated by adjusting model parameters to obtain the best visual fit of the model to time-course data obtained from TK studies. In this case, parameter fitting was accomplished through the manual process of determining one set of parameters that could consistently recapitulate a large base of diverse data. In order to minimize uncertainty in parameters, a sequential approach was followed so that the most pertinent data sets were used for each parameter. This approach to model parameterization is illustrated in Figure 3.10 and is described here.

Fitted and Calculated Parameters Time-course data on plasma MBP levels after a single i.v. dose (8 mg/kg) of the monoester were used to develop initial estimates of values

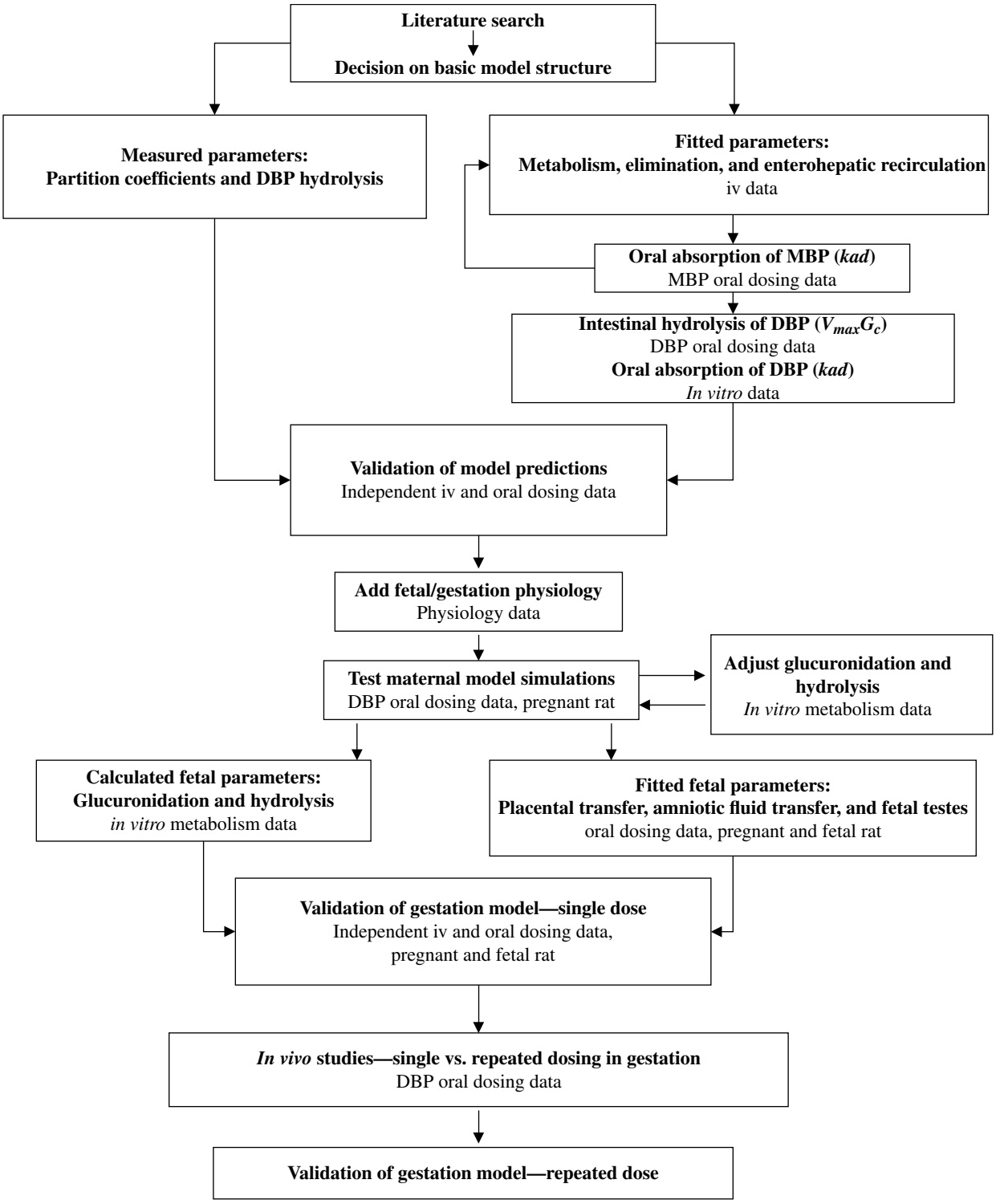


FIGURE 3.10 Model parameterization and validation for the adult male and pregnant and fetal rat. *Source:* Adapted from Clewell et al. (2008).

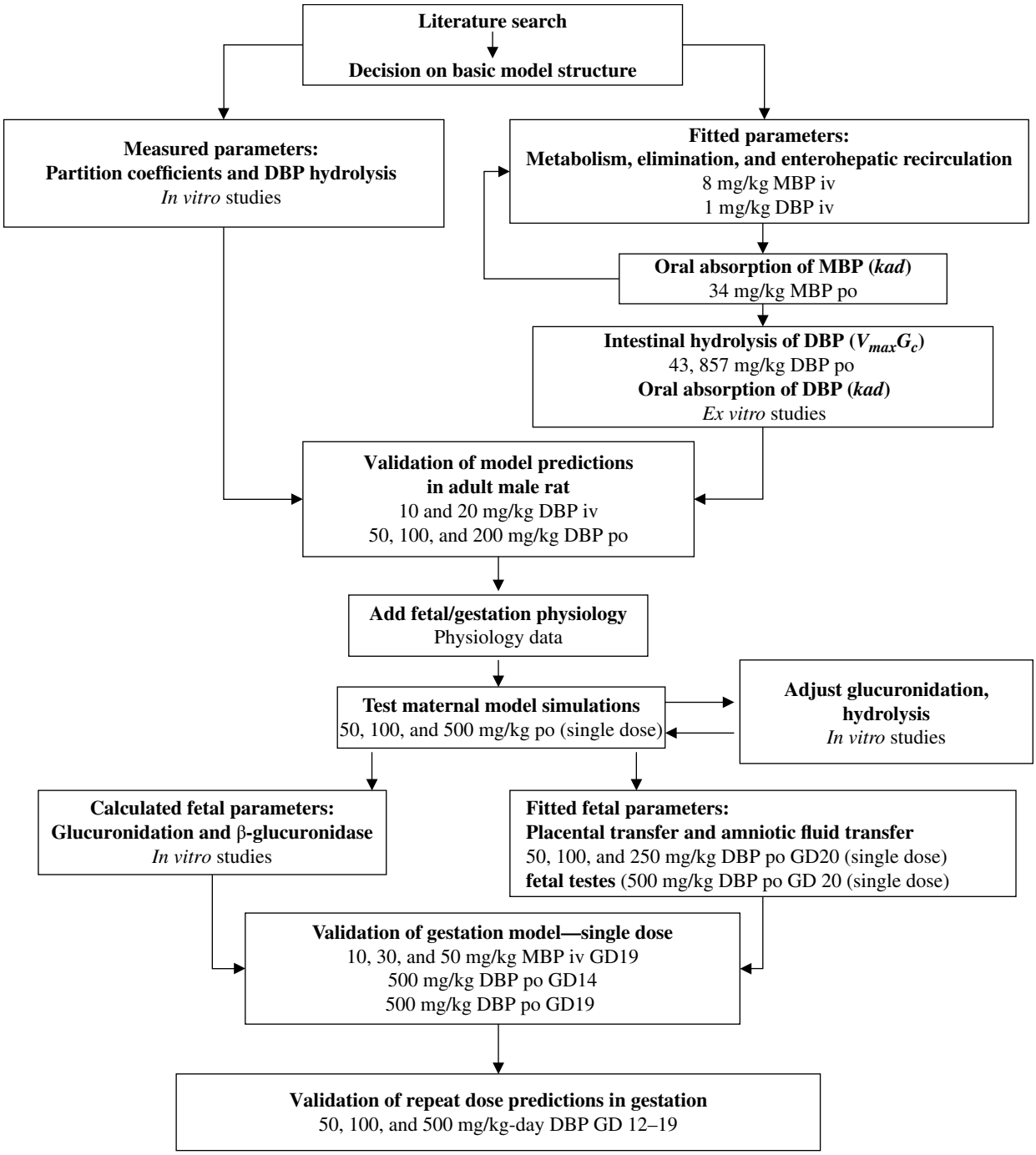


FIGURE 3.10 (*Continued*)

for the parameters governing liver metabolism and diffusion limitation in the richly and slowly perfused tissues (PARc and PASc). First approximations of the parameters governing liver metabolism and diffusion limitation in the richly and slowly perfused tissues were based on i.v. data, because it represented the simplest possible dosing scenario, and was therefore dependent upon the least number of adjustable parameters. However, as this study measured only plasma MBP concentrations, additional data were needed to refine the parameters driving elimination and enterohepatic recirculation.

A more comprehensive data set was available that utilized i.v. doses of radiolabeled DBP (14C-DBP) in the adult male rat and was ideal for estimating parameters for metabolism, elimination, and enterohepatic recirculation parameters. The i.v. study included plasma DBP, MBP, MBP-G, and total 14C

levels after a single i.v. dose of either 1 or 10 mg/kg 14C-DBP to adult male rats. Elimination of MBP, MBP-G, and total 14C was measured over 72 h in the urine. Metabolism parameters (glucuronide conjugation, oxidation), and urinary clearance rates were adjusted to fit DBP, MBP, MBP-G, and total 14C in the plasma and urine at the 1 mg/kg dose. Based on the increased water solubility of glucuronide conjugates, it was assumed for the sake of simplicity that MBP-G would be distributed with the water (VD = 0.65). The partition coefficient for MBP-G in the tissue (PGT) was then calculated from the VD, blood volume (VB), and combined tissue volume (VT) to be 0.3 (PGT = (VD − VB)/VT).

Bile cannula studies were used to set values for parameters governing enterohepatic recirculation. Bile cannulation involves inserting a shunt into the bile duct of the rat and collecting the bile before it can be transferred from the liver to the GI. Thus, chemical distribution can be compared in naive and bile cannulated animals to gauge the role of enterohepatic recirculation in the overall chemical kinetics. Bile cannulation was simulated by turning off transfer of DBP, MBP, and MBP-G from the liver to the upper intestine and biliary transport parameters were fit to the bile excretion data. The rates of MBP-G hydrolysis in the gut, fecal excretion, and movement in the gut were then adjusted to reproduce fecal excretion and plasma data in sham-operated rats.

With the earlier-mentioned parameters constrained, the first-order rate of MBP oral absorption was determined from studies in which MBP was administered via oral gavage. Finally, the parameters for DBP hydrolysis were fitted by comparing the model simulations to data from DBP oral gavage dosing, while keeping all other parameters constant.

Kinetic Parameters in the Pregnant Rat After the model was tested against validation data (data not used for parameter fitting), adjustments were made to account for the physiology and biochemical changes associated with pregnancy. When possible, the parameter values determined for the male rat were also used with pregnant rat using allometric scaling to account for differences in bodyweight. However, some biochemical changes occur during gestation that affect chemical kinetics and cannot be attributed to changes in body weight. For example, *in vitro* studies indicate that xenobiotic metabolism is altered in the pregnant rat compared to the in the virgin rat. In cases where the literature indicated that kinetic parameters would be changed, alterations to parameter values were made based on *in vitro* or *in vivo* data.

Two kinetic parameters were adjusted before using the model in the pregnant dam: the maximum capacity for DBP hydrolysis in the intestine and the maximum capacity for MBP glucuronide conjugation in the liver. Hydrolysis of DBP was decreased based on *in vitro* metabolism studies, which showed a sex difference in the ability to hydrolyze DBP, with the females metabolizing only 60% as much as the males. The maximum capacity for glucuronide conjugation in the liver was also reduced in simulations of pregnant rats. *In vitro* studies in livers of nonpregnant female and pregnant rats showed that UDPGT activity for a variety of substrates was decreased by approximately 50% during gestation (GD 19–20).

Kinetic Parameters in the Fetal Rat Parameters describing transfer between the dam and fetus and the fetus and amniotic fluid were fit to published placenta, fetal plasma, and amniotic fluid time-course data. Because no previous data was available for MBP concentrations in the placenta tissue or fetal testes, the partition coefficients for fetal tissues were fit to the available *in vivo* data.

Glucuronide conjugation was included in the fetal model based on *in vivo* data showing significant concentrations of MBP-G in fetal blood after maternal dosing and *in vitro* data, indicating that the fetal rat liver is capable of Phase 2 metabolism prior to birth. In fact, for some substrates fetal metabolic capacity is greater per milligram liver protein on GD 19 than it is in adulthood. Glucuronide metabolism of chemicals similar to MBP is characterized by low activity in the fetal liver, with a rapid increase to adult levels after birth. The parameters for fetal glucuronide conjugation of MBP were calculated from the ratio of measured fetal:adult glucuronide activities measured *in vitro*. Using the values calculated from *in vitro* studies, it was possible to predict fetal plasma MBP-G levels.

Model Validation

After setting the parameters, the gestation model was then tested with additional data sets from single-dose studies on different days of gestation (GD 14–20) and alternate dose routes (i.v., oral) (Figure 3.10). The kinetic study described previously in this chapter was used to test whether the model, which was developed based on single-dose studies, could be used to predict maternal and fetal metabolite concentrations after repeated doses. Furthermore, the kinetic studies used doses that were within the range of linear kinetics (50, 100 mg/kg/day) as well as a dose where some of the biochemical processes would be saturated (DBP hydrolysis, oral absorption, glucuronide conjugation of MBP). Some of the model predictions are shown versus validation data in Figure 3.11. The success in simulating the trend of the data across a wide range of studies (single versus repeat dosing; low versus high dose; i.v. versus po) supports the approach used to estimate parameters that are changing throughout gestation. Furthermore, the ability to predict data sets that were not used in parameter fitting suggests that the model is robust and useful for prediction of maternal and fetal tissue metabolite concentrations where data is not available.

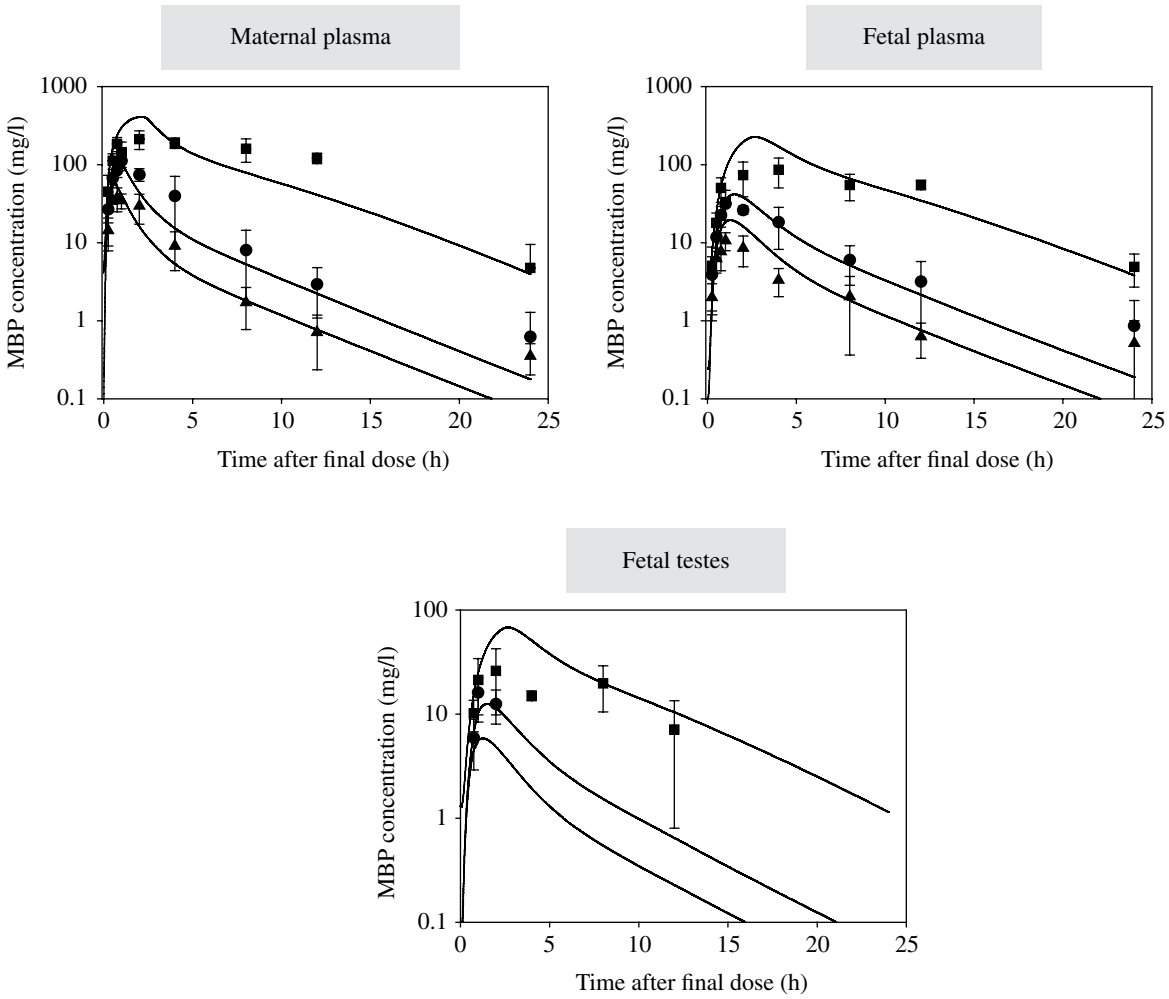


FIGURE 3.11 MBP in maternal plasma, fetal plasma, and fetal testes after the last dose of 50 (▲), 100 (●), or 500 (■) mg DBP/kg/day administered from GD 12–19. *Lines* represent model simulations. *Symbols* and *cross bars* represent the mean ± SD of the data. *Source:* Adapted from Clewell et al. (2008).

Application of PBPK Model

Several studies have shown that exposure to DBP during the period of sexual development in the fetal rat may lead to delayed development of the male reproductive tract. Dose–response examinations showed several overt effects in the male offspring from dams exposed to greater than 500 mg/kg/day from GD 12–21, including hypospadias (cleft penis), nipple retention, reduced AGD, cryptorchidism, seminiferous tubule degeneration, testis interstitial cell proliferation, malformed epididymis, seminal vesicle, vas deferens and ventral prostate. Of these effects, only nipple retention and reduced AGD have been found at lower doses (100 and 250 mg/kg/day, respectively). Delayed preputial separation has been noted in some studies at doses greater than 100 mg/kg/day, but not in others at doses up to 500 mg/kg/day. All of these adverse effects occur in concert with decreased testes testosterone concentration, which may provide a more sensitive marker for disruption of androgen-dependent development.

While these effects have been well-studied with regard to external DBP dose, there is currently no information on the internal (fetal) dose expected from these exposures. In order to illustrate how this PBPK model may be used to relate external DBP to internal MBP dose in multiple-day studies, the model was run with the repeated exposure parameters at 16 different dose levels, ranging from 1 to 550 mg/kg/day. The AUC for the last day of dosing was then used to calculate the average concentration in the maternal and fetal plasma, and fetal testes (daily AUC/24 h). The predicted dose–response is illustrated in Figure 3.12, together with reported lowest observed adverse effect levels (LOAELs) for several landmarks of male sexual development. Model predictions suggest that the average daily MBP concentration in the fetal rat testes must reach levels of approximately 1, 2, and 18 mg/l in order to cause noticeable changes in testosterone production, nipple retention, or more overt effects (i.e., reduced AGD), respectively.

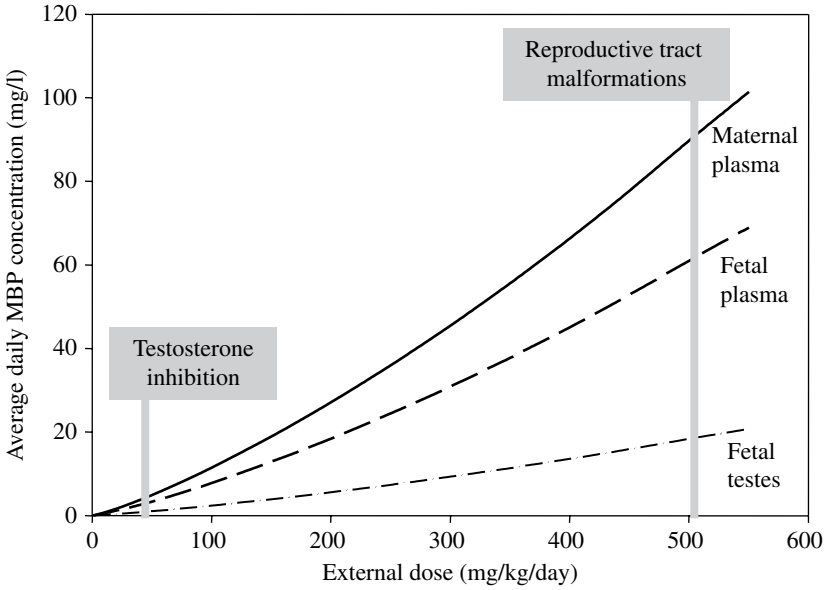


FIGURE 3.12 DBP dose–response: Predicted average daily concentration of MBP in maternal and fetal plasma and fetal testes at external doses ranging from 1 to 550 mg DBP/kg-day from GD12. *Lines* represent the average MBP concentrations in the maternal plasma (*solid*), fetal plasma (*dashed*), and fetal testes (*dash–dotted*) predicted by the model across doses. Points and drop bars represent the published LOAELs for the identified effects. *Source:* Adapted from Clewell et al. (2008).

Using the model, these testes concentrations may also be correlated to fetal and maternal plasma, which are often used as surrogates for fetal dose in the human. Maternal plasma MBP concentrations associated with reduced testosterone, nipple retention, and reproductive tract malformations in the rat fetus are 5, 11, and 60 mg/l.

3.5 TOXICOKINETICS IN THE FUTURE

Interpretation of Human Biomonitoring Data

Biomonitoring, the measurement of chemicals in human tissues and fluids, is becoming commonplace, and biomonitoring data has proved to be an important resource for identifying the presence of chemicals, both natural and synthetic, in human populations. However, the concentrations of the chemicals detected in human samples are generally very low, typically in the parts per billion (ppb) or parts per trillion (ppt) range, and the degree of risk posed by these chemicals depends on whether the exposure levels approach those known to cause toxicity in test animals or people. Unfortunately, it is often difficult to relate a measured concentration of a chemical in a human tissue or fluid to the administered doses used in animal toxicity studies.

A simple screening approach has been proposed for interpreting biomonitoring data from a risk perspective using PBPK modeling. In this approach, a biomarker concentration, referred to as the biomonitoring equivalent (BE), is determined by using the model to predict the biomarker concentration associated with continuous exposure at a given exposure guidance value or reference dose (RfD). For example, the BE blood concentration associated with the RfD for a compound would be obtained by running the PK model for continuous exposure at the RfD and setting the BE to the predicted steady-state blood concentration. The BE could then be used as a basis for determining whether a particular biomonitoring result presented a cause for concern. This approach is particularly attractive for persistent and intermediate-persistence compounds, where short-term variations in exposure have less impact. The application of the approach in the case of nonpersistent compounds is more problematic, since a random biomarker sample from an intermittent exposure could be well above or below the time-weighted average value that would be achieved by an equivalent continuous exposure, depending on the relationship of the peak exposures and the time of sampling.

As an alternative, PBPK modeling together with probabilistic analysis using Monte Carlo techniques provide the necessary tools to reconstruct exposures from biomonitoring measurements in cases where the time dependence of the exposures is likely to be important. The general approach is referred to as reverse dosimetry, that is, inferring exposure levels from the internal biomarker concentrations. This reverse dosimetry approach has been used successfully with several volatile organic compounds (VOCs), a group of compounds with complex exposure pathways and rapid clearance by metabolism and exhalation. To apply reverse dosimetry in risk assessment, the estimated distribution of population exposures can be compared with various regulatory exposure guidance values.

The two approaches, forward dosimetry and reverse dosimetry, are complementary to each other and provide different perspectives on the implications of biomonitoring data from a human health risk perspective. In either case, proper interpretation of biomonitoring data requires information on the PK and metabolism of the compound as well as information on the nature of exposures to the compound. Quantitative interpretation of human biomonitoring data can best be accomplished by linking PBPK modeling with exposure pathway modeling within a probabilistic framework.

Quantitative *In Vitro* to *In Vivo* Extrapolation

There is increasing pressure to develop *in vitro* methods for toxicity assessment in place of live animal studies. However, estimating systemic toxicity from *in vitro* results is a daunting task since it is not possible to fully reproduce the *in vivo* situation in the *in vitro* system. The absence of whole-body PK processes (ADME) in the *in vitro* assay system is one of the main challenges in implementing *in vitro*–based risk assessment for chemicals and drugs. Integration of data on the toxic mode of action of a chemical with data on its PK behaviors is essential for the interpretation of *in vitro* studies on the toxicity of a compound.

Quantitative *in vitro* to *in vivo* extrapolation (QIVIVE) is the process of estimating the environmental exposures to a chemical that could produce tissue exposures at the site of toxicity in humans equivalent to those associated with effects in an *in vitro* toxicity test (e.g., an EC_{50}, a benchmark concentration, or an interaction threshold identified by a biologically based dose–response model for the toxicity pathway of concern). QIVIVE can provide an estimate of the likelihood of harmful effects from expected environmental exposures to chemicals by integrating diverse information from targeted *in vitro* toxicity and kinetic assays using a computational modeling approach. This process is described as "quantitative" because it relates the biological responses observed *in vitro* to the expected responses *in vivo* on the basis of the equivalence of the concentrations at which the responses occur. PK modeling plays a pivotal role in this quantitative extrapolation, by incorporating the *in vivo* bioavailability, distribution, and clearance of the chemical into the process. The usefulness of physiologically based PK models in QIVIVE comes from the fact that PBPK models incorporate relevant anatomical structures that can be parameterized using independent experiments, for example, *in vitro*–derived metabolic constants, allowing them to describe the time course of the tissue concentration of a chemical at the target site for toxicity in the body. In this respect, the conversion of *in vitro*–derived PK parameters into corresponding *in vivo* values in the PBPK model is the first and also the key step for QIVIVE.

Modeling of Pharmacodynamics

The growing popularity of the PBPK modeling approach represents a movement from simpler kinetic models toward more biologically realistic descriptions of the processes that regulate disposition of chemicals in the body. To a large extent, the application of these PBPK models to study the time courses of compounds in the body is simply an integrated systems approach to understanding the biological processes that regulate the delivery of chemicals to target sites. Many PBPK models integrate information across multiple levels of organization, especially when describing interactions of compounds with molecular targets, such as reversible binding of ligands to specific receptors. In such cases, the PBPK models integrate molecular, cellular, organ level, and organism-level processes to account for the time courses of compounds, metabolites, and bound complexes within organs and tissues in the body.

The main goal of PBPK models is quite simple—to predict the target tissue dose of compounds and their metabolites at target tissues and, in some cases, to describe interactions in target tissues. PBPK models once developed are extensible. They can be used to extrapolate to various other conditions because of their biological fidelity. While the goal in applying these models is to predict dosimetry, it is important to remember that the overall goal of using PBPK modeling in toxicology and risk assessment is broader than simply estimating tissue dose, regardless of the level of detail provided in the interactions of compounds with tissue constituents. The goal in the larger context is to understand the relationship between dose delivered to target tissues and the biological sequelae of the exposure of target tissues to compounds. The specific steps that lead from these dose metrics to tissue, organ, and organism-level responses have usually been considered part of the PD process. Another inexorable development will be expansion of the systems approaches into the PD arena. This latter area will represent a systems biology approach for describing perturbations of biological systems by compounds and the exposure/dose conditions under which these perturbations become sufficiently large to pose significant health risks.

The systems biology approach (Figure 3.13) focuses on normal biological function and the perturbations associated with exposure to compounds. Perturbations of biological processes by compounds lead to either adverse responses (toxicity) or restoration of normal function to a compromised tissue (efficacy). The effects of compounds, whether for good or ill, can best be described by PBPK approaches linked to PD models of responses of cellular signaling networks. Toxicity is then defined by an intersection of compound action with the biological system. Toxicology is a discipline at the interface of chemistry/PK (primarily embedded in the vertical component) and biology/PD (primarily captured by the horizontal chain). Clearly, the main differences in the

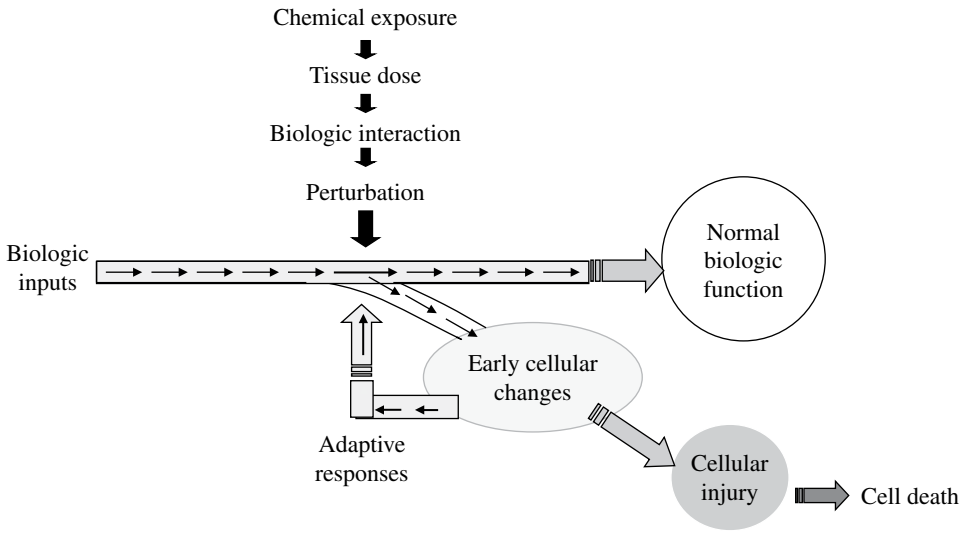


FIGURE 3.13 Modeling of cellular response. *Source:* Adapted from Krewski et al. (2010).

next generation of systems approaches in PK and PD modeling will be the increasingly detailed descriptions of biology afforded by new technologies and the expansion of modeling tools available for describing the effects of compounds on biological signaling processes.

3.6 SUMMARY

TK is the study of the time course for the absorption, distribution, metabolism, and elimination of a chemical substance in a biological system. In TK modeling, established descriptions of chemical transport and metabolism are employed to simulate observed TK in silico. Implicit in any application of TK to toxicology or risk assessment is the assumption that the toxic effects in a particular tissue can be related in some way to the concentration time course of an active form of the substance in that tissue. Moreover, in the absence of evidence for differences between species in the nature or extent of the tissue response, it is assumed that similar responses will be produced at equivalent tissue exposures regardless of species, exposure route, or experimental regimen.

TK analyses can be conducted using a wide range of model complexities, ranging from empirical noncompartmental analysis that simply describes collected data to sophisticated PBPK/PD models that incorporate a great deal of biological information. The additional effort required to develop a PBPK description is justified by the capability of such models to predict TK behaviors under conditions different from those in which the data on which the model was based were collected, including different doses, routes, and species. They also provide a platform for conducting analyses that are needed to improve risk assessment, such as quantitative IVIVE for cell-based toxicity assays, reverse dosimetry for the interpretation of human biomonitoring data, and analysis of the impact of human variability on risks in sensitive subpopulations.

REFERENCES AND SUGGESTED READING

Andersen ME. Saturable metabolism and its relation to toxicity. Crit Rev Toxicol 1981;9:105–150.

Andersen ME, Clewell HJ, Carmichael PL, Boekelheide K. Food for thought … can case study approaches speed implementation of the NRC report: "toxicity testing in the 21st century: a vision and a strategy?" ALTEX 2011;28(3):175–182.

Andersen ME, Clewell HJ, Gargas ML, Smith FA, Reitz RH. Physiologically-based pharmacokinetics and the risk assessment for methylene chloride. Toxicol Appl Pharmacol 1987;87:185–205.

Clewell HJ. Coupling of computer modeling with *in vitro* methodologies to reduce animal usage in toxicity testing. Toxicol Lett 1993;68:101–117.

Clewell HJ, Andersen ME. Use of physiologically-based pharmacokinetic modeling to investigate individual versus population risk. Toxicology 1996;111:315–329.

Clewell HJ, Gentry PR, Covington TR, Sarangapani R, Teeguarden JG.. Evaluation of the potential impact of age- and gender-specific pharmacokinetic differences on tissue dosimetry. Toxicol Sci 2004;79:381–393.

Clewell HJ, Gentry PR, Gearhart JM, Allen BC, Andersen ME. Comparison of cancer risk estimates for vinyl chloride using animal and human data with a PBPK model. Sci Total Environ 2001;274(1–3):37–66.

Clewell HJ, Tan YM, Campbell JL, Andersen ME. Quantitative interpretation of human biomonitoring data. Toxicol Appl Pharmacol 2008;231:122–133.

Clewell H, Yoon M. *In vitro-in vivo* extrapolation. In: Siegel JA, Saukko P, editors. *Encyclopedia of Toxicology*. 3rd ed. San Diego: Elsevier Inc; 2012.

Clewell RA, Clewell HJ. Development and specification of physiologically based pharmacokinetic models for use in risk assessment. Regul Toxicol Pharm 2008;50:129–143.

Clewell RA, Kremer JJ, Williams CC, Campbell JL, Andersen ME, Borghoff SJ. Tissue exposures to free and glucuronidated monobutylyphthalate in the pregnant and fetal rat following exposure to di-n-butylphthalate: evaluation with a PBPK model. Toxicol Sci 2008;103(2):241–259.

Clewell RA, Kremer JJ, Williams CC, Campbell JL, Sochaski MA, Andersen ME, Borghoff SJ. Kinetics of selected di-n-butyl phthalate metabolites and fetal testosterone following repeated and single administration in pregnant rats. Toxicology 2009; 255(1–2):80–90.

Clewell RA, Merrill EA, Gearhart JM, Robinson PJ, Sterner TR, Mattie DR, Clewell HJ. Perchlorate and radioiodide kinetics across lifestages in the human: using PBPK models to predict dosimetry and thyroid inhibition and sensitive subpopulations based on developmental stage. J Toxicol Environ Health Part A 2007;70(5):408–428.

Environmental Protection Agency (EPA). EPA request for comments on draft report on cross-species scaling factor for cancer risk assessment. Fed Reg 1992;57:24152.

Gentry PR, Hack CE, Haber L, Maier A, Clewell III HJ. An approach for the quantitative consideration of genetic polymorphism data in chemical risk assessment: examples with warfarin and parathion. Toxicol Sci 2002;70:120–139.

International Programme on Chemical Safety (IPCS). Characterization and application of physiologically based pharmacokinetic models in risk assessment. Harmonization Project Document No. 9; 2010. Geneva: World Health Organization.

Krewski D, Acosta Jr D, Andersen M, Anderson H, Bailar III JC, Boekelheide K, Brent R, Charnley G, Cheung VG, Green Jr S, Kelsey KT, Kerkvliet NI, Li AA, McCray L, Meyer O, Patterson RD, Pennie W, Scala RA, Solomon GM, Stephens M, Yager J, Zeise L. Toxicity testing in the 21st century: a vision and a strategy. J Toxicol Environ Health B Crit Rev 2010;13:51–138.

Loizou G, Spendiff M, Barton HA, Bessems J, Bois FY, d'Yvoire MB, Buist H, Clewell III HJ, Meek B, Gundert-Remy U, Goerlitz G, Schmitt W. Development of good modeling practice for physiologically based pharmacokinetic models for use in risk assessment: the first steps. Regul Toxicol Pharmacol 2008; 50:400–411.

McLanahan ED, El-Masri H, Sweeney LM, Kopylev L, Clewell HJ, Wambaugh J, Schlosser PM. Physiologically based pharmacokinetic model use in Risk assessment—Why being published is not enough. Toxicol Sci 2012;126(1):5–15.

Ramsey JC, Andersen ME. A physiologically based description of the inhalation pharmacokinetics of styrene in rats and humans. Toxicol Appl Pharmacol 1984;73(1):159–175.

Reddy MB, Yang RSH, Clewell III HJ, Andersen ME. *Physiologically Based Pharmacokinetic Modeling: Science and Applications*. Hoboken: John Wiley & Sons, Inc.; 2005.

Zhang Q, Pi J, Woods CG, Jarabek AM, Clewell HJ III, Andersen ME. Hormesis and adaptive cellular control systems. Dose-Response 2008;6(2):196–208.

4

OMICS TECHNOLOGIES IN TOXICOLOGY

Mary Jane Cunningham

Omics technologies are methods that allow global screening of all macromolecules within cells. The macromolecules include DNA, RNA, protein, and cellular metabolites. The activities of these macromolecules affect not only the cell but also all levels of the biological system from cell to tissue to organ to organism to population. "Omics" is the shortened phrasing used to describe all of the technology areas, such as genomics, proteomics, metabolomics, and pharmacogenomics. It is derived from the suffix of these terms. Each of the technology areas will be described in this chapter. It will explain:

- The origin of omics technologies
- Approaches and applications of genomics
- Approaches and applications of proteomics
- Approaches and applications of metabolomics
- Approaches and applications of pharmacogenomics
- Approaches and applications of systems biology

4.1 INTRODUCTION TO OMICS

Omics technologies have a wide range of applications. They provide scientific knowledge for better design of medical therapies by enabling screening of substances for how well they work (efficacy) or their potential toxic insult (adverse effect). We are hopeful that omics will allow scientists to examine the inner workings of the cell at a glance and thereby use this knowledge to predict interactions.

The advancements in science, particularly in medicinal and organic chemistry, have led in recent years to an abundance of new compounds for use in areas as diverse as disease diagnostics, medical therapies, chemical applications, and environmental remediation. These newly synthesized molecules undergo rigorous testing according to guidelines established and monitored by regulatory agencies. The two most common agencies are the Food and Drug Administration (FDA) and the Environmental Protection Agency (EPA). Each compound must pass a series of tests before it is given approval for the marketplace. The bottleneck of producing thousands to millions of these compounds has been addressed using high-throughput synthesis methods. However, the testing of these compounds in relevant *in vivo* and *in vitro* systems is still a hurdle. Omics technologies were developed to address this hurdle.

Another term that has surfaced is toxicogenomics. Its definition as the use of omics technologies to investigate toxicity. Toxicogenomics is a specialized branch of toxicology that uses these global screening methods to investigate how compounds act in biological systems, such as determining their mechanism(s) of action, validating the interaction with the target molecule, detailing why they may cause adverse effects, and providing clues as to how the compound can be redesigned to a less toxic compound. The information gained is useful in helping predict whether the compound will ultimately pass all safety requirements, especially before entering clinical trials with human subjects.

How Omics Came to Be

A large driver for the establishment and development of omics technologies was the field of molecular biology. In the 1960s and 1970s, molecular biology underwent huge growth as a scientific discipline. Two avenues of technological

Principles of Toxicology: Environmental and Industrial Applications, Third Edition. Edited by Stephen M. Roberts, Robert C. James and Phillip L. Williams.

advancement spearheaded this growth: (i) the development of advanced blotting techniques and (ii) the development of DNA sequencing methods. Southern, Northern, and Western blots were introduced as mainstream techniques during this period. Edward Southern developed the Southern blot to isolate and visualize individual DNA molecules. DNA fragments were isolated in an agarose gel and were visualized with intercalating dyes. The Southern blot was produced on a filter paper coming in contact with the gel and by drawing the labeled fragments out so a more quantitative and permanent record could be obtained. Shortly thereafter, blots were developed to visualize and quantify RNA (Northern blot) and protein (Western blot) molecules. While scientists were excited to finally visualize these macromolecules, these methods had limitations. They were time-consuming to perform and only a small number of molecules could be isolated and visualized at a time.

DNA sequencing was a common technique originally developed by Maxam and Gilbert. This method allowed scientists to identify the exact sequence of DNA fragments. For each fragment, the particular order of DNA bases (i.e., guanine, adenine, cytosine, and thymine) could be determined. Through later developments, significant advances were introduced by the use of terminating dyes and detection by nonradioactive labels. Therefore, DNA sequencing was made easier and more efficient without the hassle of using radioactivity.

In 1990, a national and international effort, the Human Genome Project, was started with the objective of determining the entire sequence of human DNA. The first full human DNA sequence was published in 2001 and the first phase of the Project was completed in 2003. The strategy for this project was to sequence small portions of DNA, approximately 50–800 base pairs in length, known as expressed sequence tags (ESTs). ESTs were sequenced and compiled into a public database, the GenBank. The sequences could be compared by mathematical methods, such as Basic Local Alignment Search Tool (BLAST), which allowed them to be aligned, overlapped, and compared to form a fully complete DNA sequence matching each gene. Information on ESTs formed the knowledge base for the whole human genome.

4.2 GENOMICS

The first technology area to develop was genomics. Genomics is the study of gene expression by globally screening the activity of RNA molecules. This definition, in its broadest sense, encompasses the study of messenger RNA (mRNA) expression as well as microRNA (miRNA) expression. The term refers to a variety of methods, such as polymerase chain reaction (PCR)–based assays, Northern blots, representational differences analysis (RDA), differential display (DD), rapid analysis of gene expression (RAGE), serial analysis of gene expression (SAGE), and macro- and microarrays. In its narrowest sense, it refers only to the techniques that are multiplexed and globally screening. Arrays became the prototype and are widely used and accepted. They have yielded the most quantitative and qualitative information. As with the development of all new technology fields, semantics has played a huge role. In the early literature, another term, transcriptomics, was used to refer to mRNA expression profiling but is less favored today.

To find better ways to sequence DNA and thereby study the DNA and RNA interactions of the entire cell a technology trend developed—Sequencing By Hybridization or SBH. SBH gave way to the development of global genomics screening techniques. The first SBH investigation attached molecules corresponding to ESTs with known sequences to a solid substrate. These molecules are referred to as oligonucleotides (or oligomers or oligos for short). They are a sequence of base pairs that correspond to an EST. A sequence of 30 base pairs is usually referred to as a "30-mer." The most commonly used substrate was a nitrocellulose filter. mRNA from cellular tissues was isolated and reverse-transcribed to complementary DNA (cDNA). These cDNA molecules were then labeled with detector molecules, such as ^{32}P. The labeled cDNA hybridized to the oligos on the filter. Hybridization is when a macromolecular fragment aligns and is attached to another macromolecular fragment of complementary sequence. The base pairs of both molecules line up and are noncovalently bonded. When this bonding occurs, the label is activated and the overall signal can be quantified. In this example, the detection of a radioactive signal was the indication of an active gene. Comparing control spots (from normal tissue samples) to treated spots (from samples of diseased tissues) gave a ratio of expression. If the signal for the treated spot was higher than the control spot, the activity was indicative of upregulation. If the signal for the control spot was higher than the treated spot, the activity was indicative of downregulation.

Augenlicht and coworkers were the first to publish findings using these "macroarrays." His laboratory isolated mRNA of a human colon carcinoma cell line, HT-29. The RNA was used to create a cDNA library, which was replicated onto a nitrocellulose filter in a grid format. Radiolabeled cDNA molecules were made from biopsies of patients at varying degrees of risk for colon cancer. These latter molecules were hybridized to the filters. The extent of radioactivity was scanned and analyzed enabling the activity from a wide range of genes to be detected and quantified. Two percent of the genes were either upregulated or downregulated in patients with familial adenomatous polyposis (FAP) compared to patients with low risk of colon cancer. However, the surprising result was that 20% of patients with FAP or in which the cells had not yet accumulated into adenomas were upregulated. This result suggested that gene expression changes correlated with very early stages of the cancer. A diagram depicting a typical macroarray is shown in Figure 4.1.

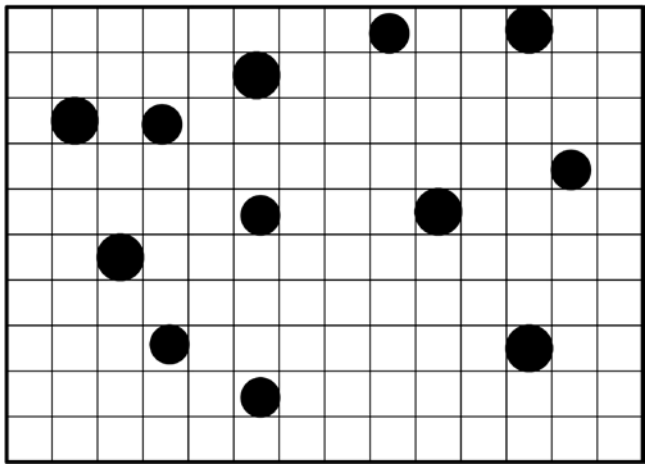


FIGURE 4.1 A diagram of a macroarray similar to that used by Augenlicht et al. The nitrocellulose filter is gridded for placement of the bacterial plasmids. Dark spots depict hybridized spotted plasmids that have increased radioactive signal.

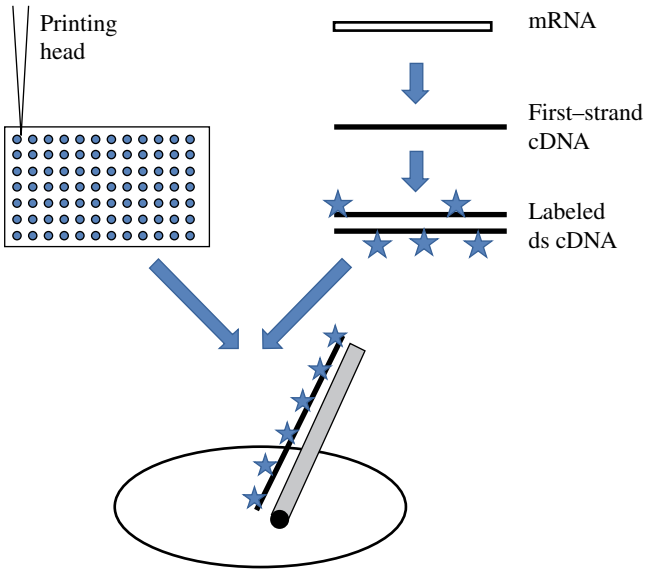


FIGURE 4.2 Microarray processing steps. Preparation of target molecules and the printing of the microarrays can occur concurrently. These steps come together at the hybridization step.

Macroarrays had limitations, however. They relied on a radioactive output. Using radioactivity meant that scientists had to adhere to strict guidelines for use and disposal, monitor all work stations, samples, and personnel, keep accurate records, and, in the experimental process, wait for the X-ray films to develop. Macroarrays also used nitrocellulose membranes, which deteriorate with age and have a short half-life. In addition, they required the use of X-ray films, which needed to be read, stored, and later disposed of safely.

To address these issues, miniaturized versions of macroarrays, "microarrays," were developed. These arrays used either silicon wafers or glass microscope slides as the solid substrate and nonradioactive labels. The microarray process involved three very basic steps: (i) printing of the microarrays, (ii) preparation of labeled molecules from cellular samples, and (iii) hybridization of the labeled molecules to the printed molecules followed by the detection and analysis of the resulting signals. These steps are depicted in Figure 4.2. The printing of the arrays and the preparation of the labeled molecules can be done concurrently. The third step of hybridization, detection, and data analysis brings the first two steps together.

Change in Semantics for Microarrays

Before going any further, it is important to note that a change in semantics occurred as the history of microarrays developed. The most widely used array formats were cDNA arrays and oligonucleotide arrays. Early SBH papers (including the citation from Augenlicht's laboratory and most of the early papers detailing work with cDNA arrays) refer to "targets" as those molecules printed onto the solid surface and "probes" as those molecules that were reverse-transcribed from RNA of the tissue samples. In the case of oligonucleotide arrays, these terms were reversed in meaning, and it is these definitions that have been adopted overall and are in use today. For clarification, the molecules printed on the solid array surface are "probes" and the molecules derived from the tissue samples, reverse-transcribed, labeled, and hybridized to the arrays are "targets."

Printing of Microarrays

As mentioned earlier, the microarrays are printed onto a solid substrate. This array printing must be done with very tight quality control measures in place and is normally performed in a clean room environment. The presence of dust or lint will interfere with the printing process. Manufacturers of arrays must make several decisions regarding what to use for (i) the solid substrate, (ii) the modification of the substrate, (iii) the type of oligo to print, (iv) modification of the oligo, and (v) the type of printing method.

Several varieties of solid substrates have been used over the years. However, the most popular are glass slides and silicon wafers. The use of silicon wafers became popular by way of the semiconductor industry. The solid surfaces must be cleaned prior to printing and this is usually done with ultrasonication or washing with acid or alkali.

The solid surface needs to be modified to ensure that the oligos attach well. One simple method is to coat the surface with polylysine derivatives. Another method involves a multistep procedure with polyethylenimine. Probably the most common protocol uses silanization. In recent years, microporous polymers are coated onto the surface so that the oligos are fixed within the gel-like substance at an orientation that optimizes their hybridization.

The next choice is what type of molecule to print. As in the case study discussed, Augenlicht's laboratory used whole bacterial plasmids containing the cDNA insert. Oligos used today are isolated and vary in length. Some manufacturers use long oligos (40-mers or longer) and others use short (25-mers and less). The overall sequence similarity and guanosine and cytosine content are important factors to consider when making this decision. It has been reported that long oligos may be optimal in sensitivity and specificity. The oligos must be checked for quality prior to printing and one way to check is to use agarose gel electrophoresis. Any product that shows more than one band or does not show any product at all is not printed and can be substituted with a "cleaner" alternative product corresponding to the same gene.

Another decision is whether to modify the printed molecule. Some investigators feel that a further modification may help with array stability. Some modifications used have been the addition of an amino group at either at the 3′- or 5′-ends or using thiol, disulfide, or benzaldehyde groups.

Finally, it must be decided how to perform the printing of the arrays. These methods have been divided into contact and noncontact methods. Contact methods use microspotting devices to deliver the oligo solution to the printing surface. These devices deliver the solution by capillary action and tips are either disposed of or washed thoroughly before the next printing. Initial devices were hand-built using an *XYZ*-axis gantry robot with the macromolecular solutions being delivered through stainless-steel printing pins. Various pin heads are now available in materials that are easier to clean, such as ceramic and tungsten. Noncontact methods disperse the oligo solutions onto the array surface without the dispensing tool touching the array surface. Many of these devices use piezoelectric technology where an electrical current controls droplet formation to precise measurements. Once the oligos are printed, the surface is treated to ensure a tight covalent bond and to lessen nonspecific binding during hybridization. Common procedures crosslink the oligomers to the surface with ultraviolet (UV) light, bake the slides in a vacuum oven, or treat with succinic anhydride or a mixture of succinic anhydride and acetic anhydride.

Finally, once the oligos are printed on the arrays, it is necessary to obtain an accurate quantification and to assess their quality. A simple tool is microscopic examination of the slides with food coloring. Another tool is to pull arrays using random sampling and hybridize these samples with a labeled PCR primer (usually one used in making the oligos). Keep in mind that this hybridization will destroy the active sites on the probes. Nondestructive means have successfully employed dyes such as dCTP-Cy3 or SYBR Green.

Preparing Target Molecules

Target molecules represent the molecules in the cell or tissue under study. The first step is to isolate RNA from the cells or tissues using a buffer containing guanidine isothiocyanate or a derivative. For tissues, this step is usually performed with homogenization. As a note of caution, tissues can vary widely in their characteristics. Modifications to the standard protocol may be necessary if the tissue is, for example, fibrous or fatty. The cell-buffer solution is then extracted with phenol/chloroform, and total RNA is purified using specialized centrifuge columns, which allow binding of the RNA. The RNA is eluted afterward and quanitified. At this point, a sample may be run on an agarose gel or added to a microchip and analyzed with a device that gives a quantity and quality readout. One such device is the Bioanalyzer manufactured by Agilent Technologies.

The total RNA is then used as a template to synthesize a labeled cDNA molecule. The first reaction (reverse transcription) synthesizes a strand of DNA from the RNA template. The second reaction synthesizes the other strand of DNA. The finished product is a double-stranded cDNA, which is subsequently labeled with a detector molecule, such as biotin.

The preparation of the target molecules may differ depending on which type of microarray is used. Some manufacturers now make reverse-transcribed cRNA molecules. Others require the labeling to be at a different point in the process.

Hybridization of Microarrays

An aliquot of the labeled cDNA is added to the microarray, and the hybridization process is allowed to proceed. The hybridization is most commonly done for 16–20 h. The arrays are then washed with buffers of varying stringency to ensure that only the hybridized pairs remain on the slide and all unattached molecules are washed off. The arrays are then scanned at the wavelength of the detector molecule, and the signal is quantified. The steps involving scanning the arrays and detecting the signal and quantifying it are referred to as "image analysis." The analysis that takes place after these steps is referred to as "data analysis." In the literature, these two types of analyses are often mistakenly used interchangeably.

Data Analysis for Gene Expression Microarrays

Analysis of microarray data varies from investigator to investigator and in complexity. Generally, initial analysis of the data is done by looking at it globally and then more specifically for the activity of particular genes. However, the first step that may be required is filtering and normalization. The data obtained from the image analysis step must be filtered for background signal and signal quality. Then, depending on the experimental design and the technology platform, the data must be normalized so that it can be compared on the same scale. The scanner detects the signal on each array and the signals must be normalized within an array as well as between arrays. The remaining steps will

analyze "expression profiles." These profiles are all the signals detected for all the genes for a particular sample or they can be all the signals detected for a particular gene for all the samples used in the experiment. Expression profiles can vary by time, dose, or other experimental conditions.

The next step is to analyze the filtered and normalized data with mathematical algorithms to get a global picture of the activity of the genes. Two common methods are hierarchical agglomerative and K-means clustering. These methods compare the expression profiles over all the genes in the experiment or over all the samples in the experiment. Slightly different results can be obtained by varying the different distance and linkage metrics or parameters in each method. It is the way comparisons are made in the different methods that makes them different. Hierarchical agglomerative clustering groups together the data in a stepwise fashion matching one expression profile with the next, one at a time, until all the profiles are matched and form clusters, as shown in Figure 4.3. K-means clustering starts out by randomly clustering the expression profiles into groups or bins and then performing the comparisons between the bins, by selecting like profiles to stay and putting different profiles into other bins, until all the profiles are matched and reorganized into bins or clusters that show similarity. The most common way to visualize the data is using "heat maps" (Figure 4.4). These maps are grids of usually two colors, such as red and blue, with rows representing the expression signals across samples and columns representing expression signals across genes. Black usually represents a signal in the treated sample the same as the signal in the control sample or a nonsignificant gene expression. Red (or in the figure, light gray) usually represents a signal higher in the treated sample compared to the signal in the control and signifies "upregulation." Blue (or in the figure, dark gray) usually represents a signal lower in the treated sample compared to the signal in the control sample and signifies "downregulation."

Other methods are used to analyze the data to pull out significant trends. For example, investigators may use higher-level mathematics to identify upregulated or downregulated genes across the experimental parameters. These genes, or biomarkers, may correlate with the disease condition being investigated or the pharmaceutical agent being studied. Biomarkers are molecules whose activity may signify a high correlation with a disease or condition or exposure to a particular agent. Indeed, comparisons are routinely made of upregulated and downregulated genes in cellular pathways that may provide clues as to adverse effects or mechanisms of action. Biomarkers can be any gene, protein, or metabolite and will be referred to later in these contexts.

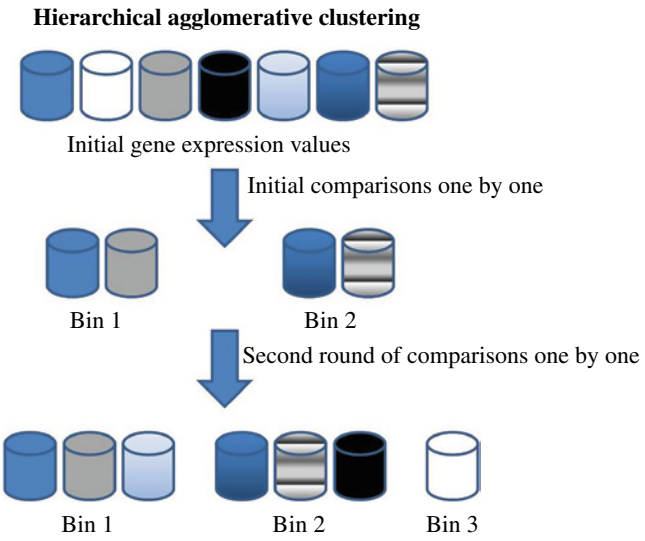


FIGURE 4.3 Hierarchical clustering binning.

Microarray Platforms

Four main microarray technology platforms were developed: (i) oligonucleotide arrays, (ii) cDNA arrays, (iii) electrokinetic arrays, and (iv) fiberoptic arrays (Figure 4.5). The first two types are the most common platforms in use today. However, two other platforms are gaining popularity: a modification of the electrokinetic platform and a solution-based platform.

Oligonucleotide arrays and cDNA arrays were developed in the same time period. Oligonucleotide arrays were manufactured using photolithography. This process builds the macromolecular sequence *in situ* or at a specific location on the solid substrate surface. A mask is put over the array surface to allow chemical reactions to take place only at specified sites. The reactions involve photolabile groups activated by light to free hydroxyl groups. The hydroxyl groups are then covalently bound with phosphoramidite-activated deoxynucleosides. The bound molecule is capped and oxidized. After rinsing, a new mask covers the array surface, and a new set of reactions occur. With each round of reactions, an activated nucleoside is added and the oligonucleotide (oligo) is synthesized to approximately 25 nucleotide oligomers or 25-mers.

Later, two other methods were developed. One method uses a digital micromirror to create a virtual mask for the photolithographic manufacturing process. The other method allows parallel synthesizing of several different types of molecules. Only the former is still in use as a marketable product today.

cDNA arrays were made by printing oligos onto a glass slide. Initially, the length of the oligos varied between 40 and 100s of base pairs. Today, most commercially available microarrays use short- to medium-length oligos as the molecules of choice with the actual length varying between manufacturers. In addition, most manufacturers now use a modified surface for printing so that the orientation of the oligos is optimal for hybridization.

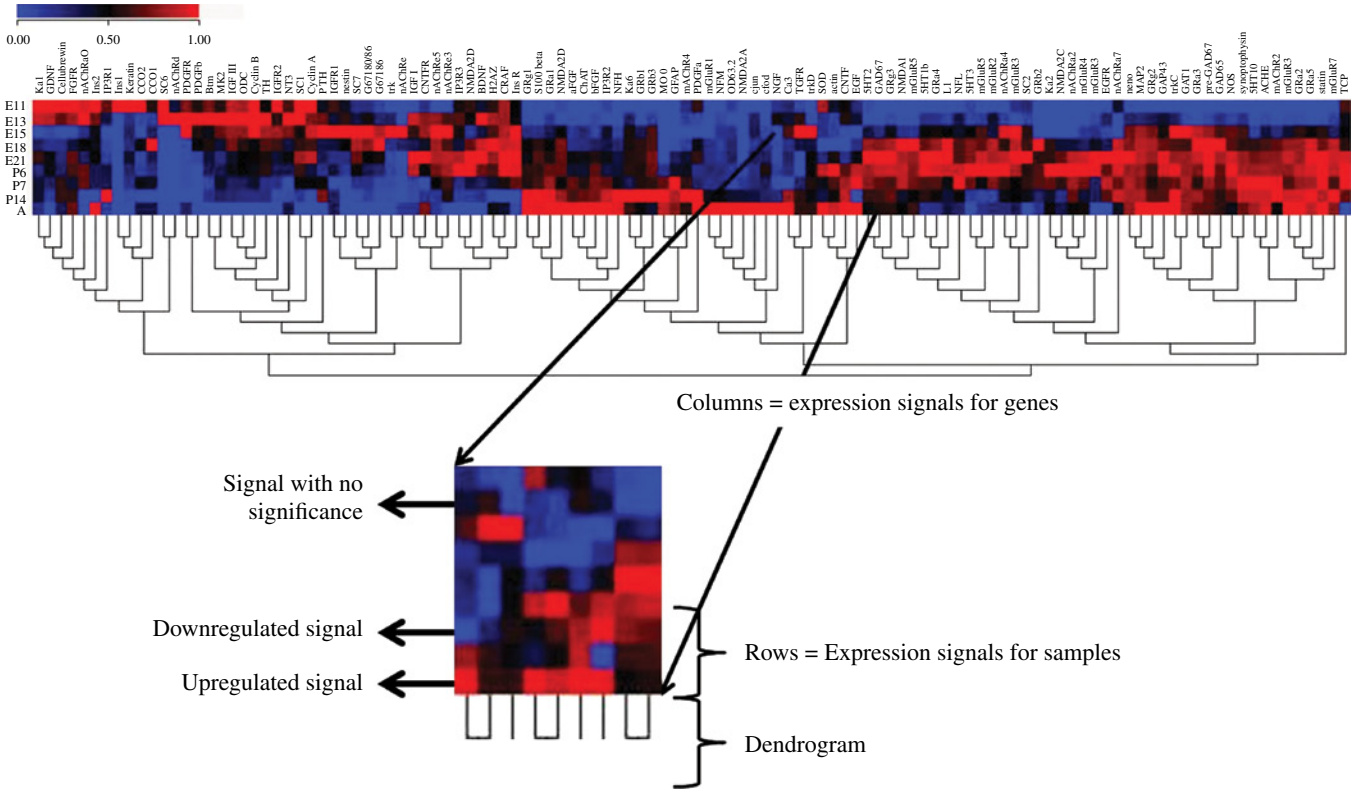


FIGURE 4.4 A heat map showing clustering results. Rows depict expression profiles of samples and columns depict signals corresponding to genes. Light gray, upregulation; Dark gray, downregulation; black, no significant gene expression.

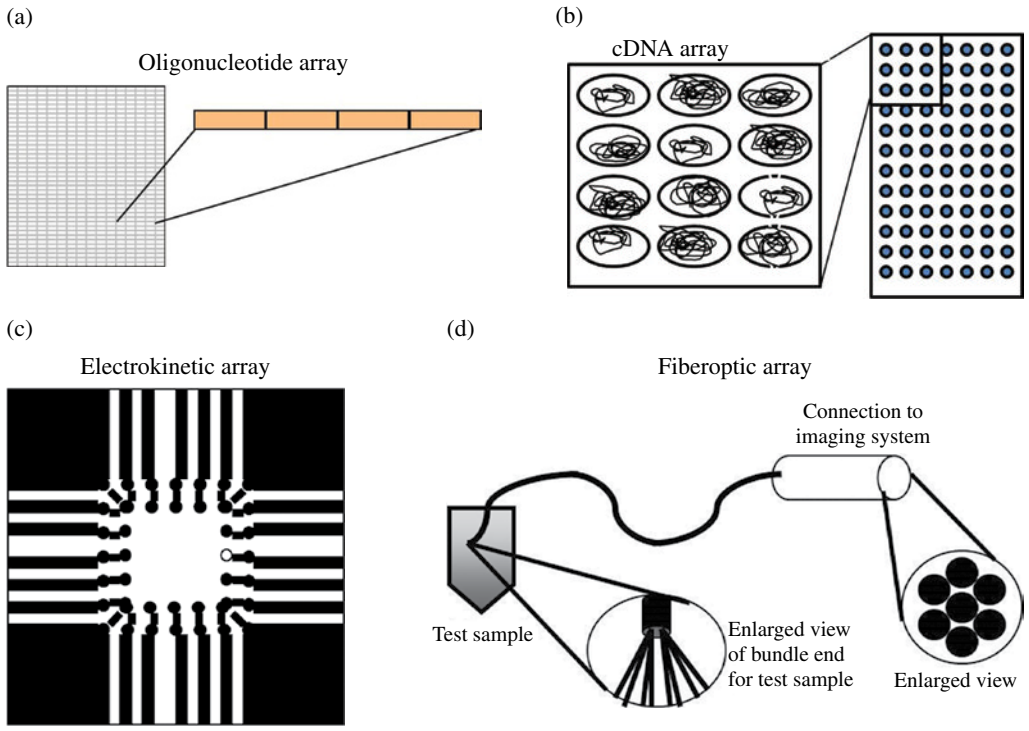


FIGURE 4.5 Depictions showing all four microarray formats: (a) oligonucleotide, (b) cDNA, (c) electrokinetic, and (d) fiberoptic.

Electrokinetic and fiberoptic arrays were early alternatives incorporating an electrical signal. Electrokinetic arrays were made with test tiles on the array that included a cDNA molecule. The hybridization of target molecules to these test tiles occurred and could be reversed by turning an electric current on and off. Fiberoptic arrays had the probes attached to an end of a fiberoptic cable. The hybridization took place by dipping the cable into a solution of target molecules, and the rate of this reaction was detected by reading the output of an electrical current.

Initially, a significant factor played a role in the distinction between oligonucleotide and cDNA microarrays. Formerly, oligonucleotide arrays were referred to as "one-color." They used one label for the target molecule, and one target molecule corresponding to a cellular or tissue sample was hybridized to an array. cDNA microarrays were referred to as "two-color." They used different target molecules for the treated tissue sample and the control sample. Usually, the target molecule for the control sample was labeled with a green dye detector molecule and the target molecule for the treated sample was labeled with a red dye detector molecule. At the hybridization step, equal amounts of the red- and green-labeled target molecules were cohybridized to the probes. It was a competition reaction based on numbers. The arrays were then scanned at ranges for both the red and green dyes and the signals compared as a ratio.

The design of the overall experiment depended on whether a one-color or a two-color system was being used. A one-color system requires that the control and treated samples are run on separate arrays, and therefore intraarray as well as interarray variability needs to be considered and accounted for. A two-color system requires that both the control target and the treated target compete for a complementary probe molecule on the same array. In this system, the importance of the intraarray variability is lessened.

The technology platform of microarrays is continuing to be developed and broadened to include other molecules besides DNA and RNA. Platforms have been developed to incorporate proteins, peptides, chemicals, tissues, and cells. Protein microarrays will be discussed in the next section. Tissue arrays have been made by microtoming a tissue into slices and adhering these slices to a solid substrate. Most commonly, lectins and antibodies used in classical pathology are immunoprecipitated onto the slices to detect their presence in the tissue. In this way, an entire tissue can be sectioned and analyzed instead of a random sampling of slices. Another platform is composed of entire individual cells adhered to a solid substrate in order to fully investigate the molecular activity per a particular cell.

MicroRNA Expression Profiling

Gene expression microarrays have traditionally been set up to monitor the activity of genes. However, in reality, they monitor this activity indirectly. The initial total RNA preparation used to make the target molecules contains all of the species

of RNA: mRNA, transfer RNA (tRNA), and ribosomal RNA (rRNA). However, it is the mRNA whose activity is in reality monitored.

Recently, another species of RNA has been discovered, miRNA. miRNAs are endogenous, noncoding small RNAs of about 21–25 nucleotides in length. They are one of the most abundant classes of gene regulatory molecules in plants, animals, and viruses. The first miRNA reported, *lin-4*, is a gene that was known to control the timing of larval development in *Caenorhabditis elegans.* It was found not to code for a protein but instead produced a pair of small RNAs. This miRNA was found to substantially reduce the amount of LIN-14 protein without noticeable change in levels of lin-14 mRNA.

Subsequently, miRNAs were found to regulate gene expression by targeting mRNAs for cleavage or translational repression. The target mRNA is silenced and so is the specific gene expression. Therefore, when a miRNA is upregulated, it is more likely that the target mRNAs are silenced. When a miRNA is downregulated, the expression of the target mRNAs is enhanced. The only caveat is that this relationship is not a one-to-one relationship. Several miRNAs may bind one mRNA, working cooperatively to silence gene expression. On the other hand, a single miRNA can affect the expression of a large set of mRNAs.

The mechanism of how miRNAs work has been covered in more detail in various review articles. Briefly, a primary transcript (pri-miRNA), of a length reaching and possibly exceeding 1 kb, is transcribed from DNA and is cleaved by an RNase, Drosha, working in combination with a double-stranded RNA-binding domain protein, DGCR8, into a pre-miRNA. The pre-miRNA is transported out of the nucleolus to the cytoplasm where it is cleaved by a second RNase, Dicer, its double-stranded structure unwound by a helicase, and is then incorporated into the RNA-induced silencing complex, RISC. The complex helps the single-stranded miRNA bind to the target mRNA molecule. These cellular mechanisms are depicted in Figure 4.6.

Initially, investigators studying miRNA expression profiling used mRNA gene expression microarrays. They searched for missing signal suggesting the transcript had been cleaved or repressed. In recent years, a database of miRNA sequences has been developed. From this knowledge base of information, several manufacturers have developed miRNA gene expression arrays

4.3 PROTEOMICS

Proteomics refers to the study of protein expression. These technology platforms apply high-screening methods to identify protein signatures within cells and tissues. They are the protein counterparts to genomics. These methods allow the investigator to simultaneously look at the whole complement of proteins belonging to a cell or tissue. Proteomic technology platforms developed from earlier applications of polyacrylamide gel electrophoresis (PAGE) and Western blots. Early studies separated proteins using two-dimensional electrophoresis (2DE) by size and pI. Cells or tissues were lysed with detergents and then stabilized. The lysates were separated by pI or charge using isoelectric focusing. Each tubular gel was then laid across one end of a slab 2DE polyarylamide gel, and the proteins were separated by mass. Silver stain was then used to detect the resulting patterns of protein spots or features. The spots were excised and analyzed by mass spectroscopy (MS). The protein corresponding to a feature was identified in two different ways: (i) by the mass-to-charge ionization pattern and (ii) by tryptic digestion and sequencing. In the first way, the molecules are ionized and give a distinct pattern based on their mass-to-charge ratios. In the second method, the protein features were excised from the gels, digested with trypsin, and analyzed by MS. This MS method analyzed the sequence of amino acids within the peptide fragments. The final data containing the mass, pI, ionization pattern, and the amino acid sequence of the peptides was compared to data from known samples to identify the protein correctly. This traditional platform of proteomics is depicted in Figure 4.7.

Investigators have attempted over the years to correlate protein expression profiles with gene expression profiles to piece together the interactions of RNA and protein within cells. Two different studies, one using rat liver and one using yeast, showed that the correlation between RNA and protein abundance was not a 1:1 ratio. In fact, the ratio in rat liver was 0.48 and the ratio in yeast ranged between 0.1–0.4 and 0.94 depending on whether the proteins were in the low-abundance class or the high-abundance class. The explanation for a less than 100% correlation between RNA and protein may be posttranslational modifications (PTM) of proteins or degradation of mRNAs preventing the full synthesis of the mature proteins.

Proteomic methods do have their limitations. More than one protein can be present in a feature. This complexity can cause confusion in being able to identify the protein composition and can lead to a failure to detect an accurate quantity. In addition, the gels deteriorate with age and can only be sampled for a short period of time. Also, the gels need to be made and run under very tight quality assurance parameters to avoid distortion and so that they can be matched to each other. One protein lysate is added to one gel, but the gels are compared across an experiment. Most investigators use control protein, and gel markers to address this issue. Finally, PTM can be detected as a series of spots running across the gel from the main protein entity. While this observation can help identify the presence of PTM, these series of features often appear as blurred and streaked leading to difficulty in their quantification.

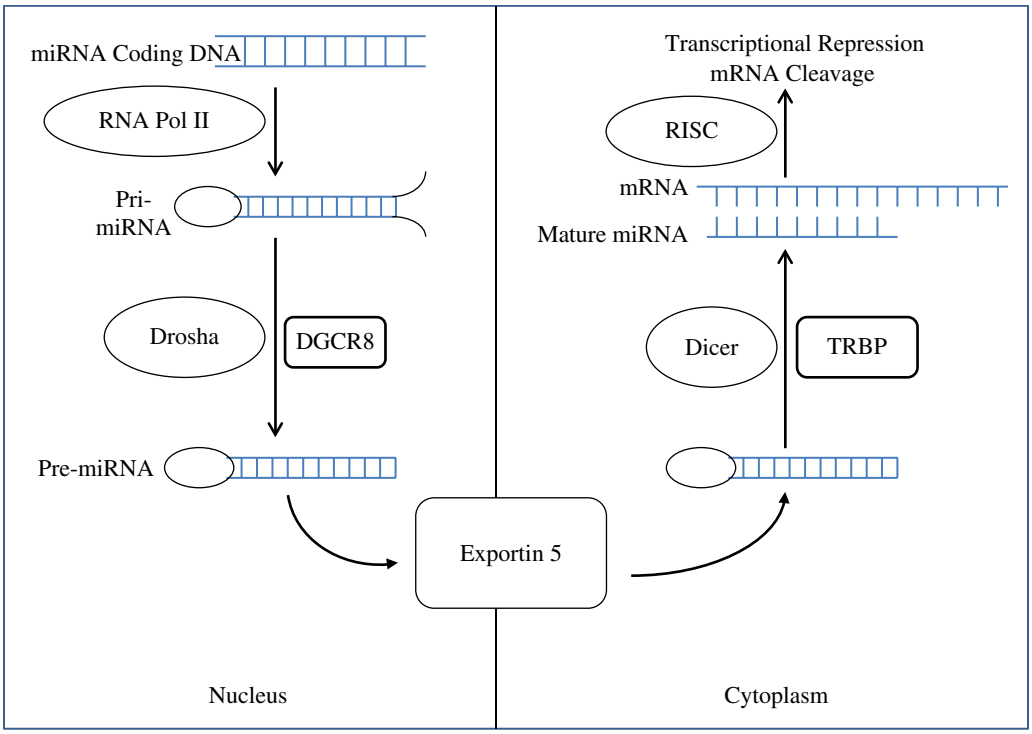


FIGURE 4.6 A diagram depicting miRNA biogenesis. RNA polymerase II transcribes miRNA genes generating primary miRNA transcripts (pri-miRNAs). Dosha and its companion protein, DGRB8, then process the pri-miRNAs into the approximately 70-nucleotide hairpin precursor miRNAs (pre-miRNAs). These pre-miRNAs are exported into the cytoplasm by Exportin 5 where they are further processed into unstable duplex structures by Dicer and its associated protein, TRBP. The duplex guide strand is then incorporated into RISC, which regulates protein expression.

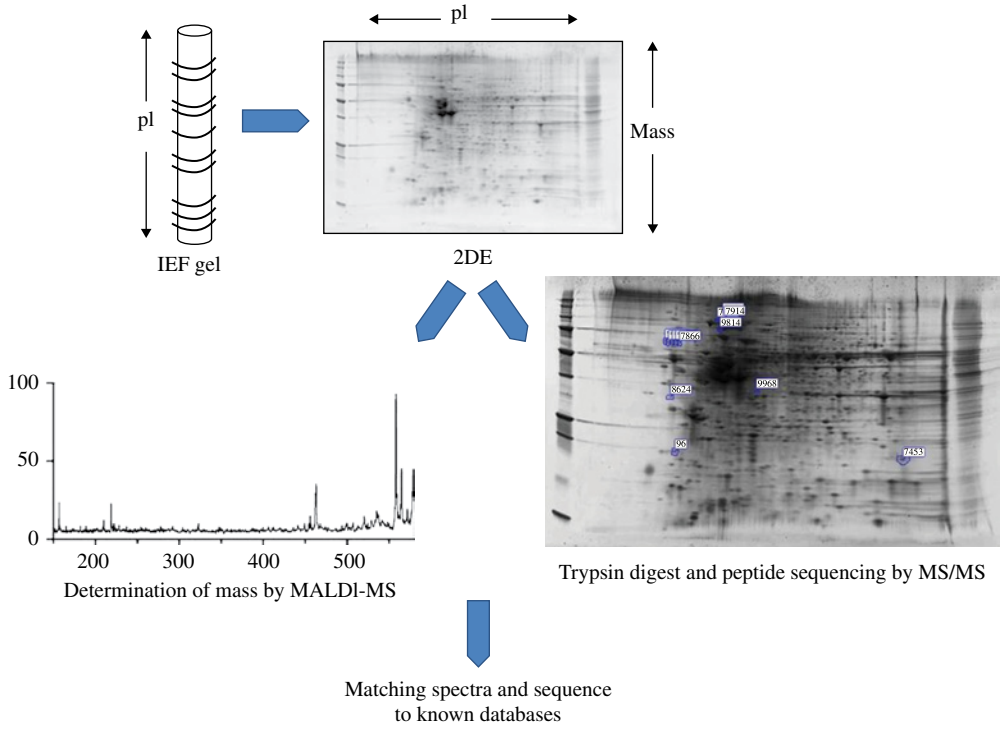


FIGURE 4.7 Traditional proteomics technology platform.

Several newer proteomics technologies have come into existence since these early studies. Investigators have printed proteins onto solid substrates to make protein microarrays. The hybridization reaction is really an antigen–antibody interaction with the target molecules being monoclonal antibodies. These protein microarrays will identify the presence of antigens in a cellular sample in much the same way as gene expression microarrays identify mRNA species. The most recent advancement is the development of label-free quantitative mass spectrometry (LFQMS). For this method, the proteins are first digested with a protease into a peptide mixture. The mixture is analyzed by liquid chromatography–MS (LC–MS) or LC–MS/MS and subsequently identified by database searching. It is rapid, sensitive, scalable, and has an increased dynamic range compared to the traditional 2DE method.

4.4 METABOLOMICS

Metabolomics is the expression profiling of metabolic changes in a biological system. One will see the phrases "metabolomics" and "metabonomics" used interchangeably. Metabolomics was initially referred to as metabolic changes within a cell, while metabonomics was the metabolic changes within a biological system. Several investigators refer to this field as "global metabolite profiling," rather than pick which word to use. In either case, the focus is on tracking the changes in "the end products of cellular regulatory processes." The ultimate goal is to be able to profile simultaneously and over time what the small molecules within the cells are doing and what roles they are playing.

To be able to obtain this goal, equipment and instrumentation previously used in standard metabolism studies have been updated and revolutionized to be able to monitor and detect these end products or metabolites in a global fashion. Previously, metabolism studies used various spectroscopy or chromatography tools. Initial studies relied on high-pressure liquid chromatography (HPLC). The development of metabolomics has forced these technologies to become more sophisticated. Nuclear magnetic resonance (NMR) spectroscopy was advanced by increasing the magnetic field strength to improve spectral resolution and sensitivity factors. Chromatography methods have been improved by decreasing the diameter of the column packing material (thereby increasing the interactive surface area) and developing columns that can withstand higher operating pressures, such as HPLC or a newer tool, ultraperformance liquid chromatography (UPLC). Another development is in using these tools in tandem or coupling the different capabilities. For example, early investigators used GC–MS or LC–MS for metabolomic studies. Now, investigators use UPLC–MS, HPLC–MS, or even HPLC–NMR–MS. Finally, chip technologies are being applied in the sample preparation step before MS.

The first metabolomics studies were investigations of metabolite patterns in complex biological fluids. Early studies reported using proton NMR to analyze metabolic profiles in blood, plasma, urine, and cerebrospinal fluid. Toxicity in liver and kidney has been investigated by monitoring the metabolite patterns in rats exposed to carbon tetrachloride, a-naphthylisothiocyanate, 2-bromoethylamine, and 4-aminophenol. All were known toxicants, and adverse effects were expected. However, the profiles obtained with metabolomics showed differences at detection levels lower than with traditional methods, such as histopathology or clinical chemistry. More recent analyses have included monitoring metabolite profiling in tissues and cell culture.

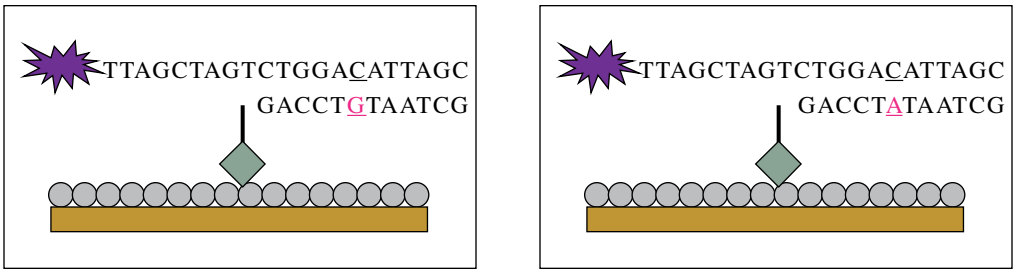
What are the benefits of using metabolomics technologies? If one can profile metabolites produced from the breakdown of a drug or produced by the cell after an interaction with a substance, one can further determine if these actions lead to any adverse effects. Drugs (and other compounds) are metabolized by cellular enzymes that break down these compounds into (i) molecules that interact with cellular entities or (ii) molecules that are modified and more readily excreted or disposed of. It was previously difficult to sift out which metabolites were necessary and needed for a therapeutic result and which were not. Metabolomics allows for simultaneous monitoring of these metabolites and may ultimately be used in real-time detection of therapies or toxicities as they happen.

4.5 PHARMACOGENOMICS

Pharmacogenomics is the use of high-throughput screening techniques for the detection of an individual's genetic variation and its role in efficacy and safety. Pharmacogenomics and pharmacogenetics are two terms that have been used interchangeably and may be defined differently according to some scientists. Pharmacogenetics is a field of study that has been around for several decades. It has been defined as the study of an individual's response to a drug as determined by their genetic makeup. Early recorded observations focused on patient variability in metabolizing various drugs, such as succinylcholine, dibucaine, and primaquine. In 1956, Werner Kalow, a pioneer in this field, first noticed variability in patient response that correlated with differences in pseudocholinesterase level. In this section, pharmagenomics will be the term used and is defined as using expression profiling to discover the variations in an individual's genotype, which explains the differences in phenotypic response. The expression profiling focuses on the detection of single-nucleotide polymorphisms or SNPs.

An example of a SNP is shown in Figure 4.8 where the most widely occurring sequence is the wild type and a genetic variation is depicted as the variant.

Most studies in pharmacogenomics have focused on variations within the drug metabolizing enzymes (DMEs). These



Perfect match Mismatch

FIGURE 4.8 Microarrays of SNPs. This illustration depicts a perfect match sequence corresponding to the wild type and a mismatch sequence corresponding to a variant.

DMEs catalyze the Phase I and Phase II metabolic reactions. Variations in genetic sequence leading to different drug responses have been observed for: arylamine *N*-acetyltransferase (NAT1 and NAT2), thiopurine-*S*-methyltransferase (TPMT), glutathione-*S*-transferase (GST), as well as the various cytochrome P450 enzymes, such as CYP2A6, CYP2C8, CYP2C9, CYP2C19, CYP2D6, and CYP3A4. Early studies showed individuals with these genetic variations can be classified as "poor," "extensive," and "ultrarapid" metabolizers. This variability is extremely important because poor metabolizers could accumulate the drug within their tissues causing a toxic response, while extensive metabolizers break down the drug so rapidly that it may not stay around the tissues long enough to exert its desired effect. For example, an individual with the gene variant CYP2D6*2, which represents an "ultrarapid metabolizer," can possess a much higher activity for cytochrome P450 isozyme 2D6, whereas an individual with the gene variant CYP2D6*10 possesses a defective enzyme and is therefore a "poor metabolizer." Individuals who have slow or fast metabolism rates have been observed in several other mammalian species as well, such as rats, mice, monkeys, dogs, and rabbits.

SNPs were detected in a high-throughput manner using oligonucleotide microarrays that contained perfect match and mismatch sequences. The changes in signal detected between the perfect and mismatch sequences indicated a presence of a SNP. The perfect match is the sequence of the wild type and the mismatch is the sequence of the variant. HIV, breast cancer, and cystic fibrosis are among the diseases studies using SNP microarrays. Recently, an array containing over 11,000 SNPs was used to detect genetic variabilities in over 100 DNA samples. This study illustrated a high-throughput variation on association studies.

4.6 SYSTEMS BIOLOGY

Systems biology is investigating an entire biological system from the molecular, cellular, tissue, organism, population, and, finally, ecosystem levels. It applies comprehensive computational and network biology methods to analyze the data produced by the omics technology platforms. The ultimate goal is to predict compound-cell and cell–cell interactions in silico.

In order to study systems biology, an experimental design should be used that not only answers the questions being asked but includes an emphasis on validation. In other words, the design should include some measures of repeatability so that checks and balances are incorporated and used to verify the technology platform. Particular attention needs to be paid to (i) types of species, (ii) gender, (iii) cell or tissue system, (iv) treatment scheduling, dose, and route, (v) appropriate controls, (vi) numbers of replicates, (vii) correlation with independent morphological and pathological toxicity assays, and so on. These choices should be able to minimize the variability in the experimental process so that the biological variability is not hidden and can easily be detected. In the case of gene expression microarrays, the Microarray Gene Expression Data (MGED) Society has proposed a list of guidelines for microarray work, known as the Minimum Information About a Microarray Experiment (MIAME) guidelines. These guidelines are recommended requirements for how experiments are designed and reported.

To obtain the goal of being able to predict compound-cell and cell–cell interactions using omics technologies, reported studies have used the same samples on different technology platforms. The combined data sets are analyzed to identify the effects on all molecules: genes, proteins, and metabolites. Early studies attempted to correlate genomics with proteomics systems (as discussed earlier) and metabolomics with genomics. A fully integrated approach using genomics, proteomics, and metabolomics was attempted by two groups of investigators. Schnackenberg and coworkers studied the hepatotoxicity effects of valproic acid. Correlative changes in glucose metabolism were shown by both metabolomics and proteomics methods. However, no significant changes were observed with the gene expression method. Kleno and coworkers used genomics, proteomics, and metabolomics to investigate the hepatotoxicity of hydrazine. Disruption in both glucose and lipid metabolism were present in all three data sets and therefore, once again, these three technologies were shown to be complementary.

Trends in Data Analysis

Expression platforms generate a large amount of data and therefore necessitate using analysis methods that can handle on the order of 10^6 to 10^{10} data points and reduce it to a 2D or 3D format for easy visualization. Each method used today takes a unique approach to analyzing the data and the method chosen may depend on the questions being asked. Some of these methods aid in reducing the large amount of data to a more reasonable data set from which more directed queries are made. Other methods may allow a more directed approach to identify outliers in the data set. These outliers may prove to be potential drug targets or markers of efficacy or toxicity. The ultimate goal is to evaluate the data in the most meaningful manner and to be able to visualize trends for further study.

Most methods are separated into two categories: unsupervised and supervised methods. Unsupervised methods include hierarchical agglomerative clustering, K-means clustering, self-organizing maps (SOM), support vector machines (SVM), and principal components analysis (PCA). These algorithms answer the question of how similar or dissimilar the data is and can be applied to all omics technology platforms. The clustering methods were previously covered under genomics. Supervised methods include algorithms based on Bayes theory and/or artificial neural networks. These procedures require the user to first select a random sampling of the data. The algorithm uses this sample to match the data with a particular class or parameter of the experiment and become "trained." Then the algorithm is applied to the entire data set and the correlations are made. The methods incorporate statistical tools to help clarify when the algorithm was successful or not. The protocols help identify outliers that correspond to active genes, proteins, and metabolites or biomarkers.

Methods for visualizing the data are numerous, but the end objective is to summarize the large data sets as trends. The trends themselves may not be conclusive but can provide valuable information to refine hypotheses and further analyze subsets of data.

4.7 SUMMARY

Advances in molecular biology led to the development of "omics" technologies. These technologies allow global screening of genes (genomics), proteins (proteomics), metabolites (metabolomics), and SNPs (pharmacogenomics). Microarrays are most widely used in genomics and enable the expression of tens of thousands of genes to be simultaneously analyzed. Arrays have also been applied to studies of miRNA and protein expression as well as SNP detection. The ability to monitor the whole complement of genes, proteins, and metabolites within an entire cell or tissue has spurred the development of sophisticated data analysis methods. Trying to understand patterns by analyzing millions of data points makes the reality of understanding what is happening at the molecular level within a cell or tissue possible. In toxicology, the application of omics technologies and systems biology sets the stage for rapidly providing answers to hazard assessment and adverse effects of compounds.

SUGGESTED READING

Adams MD, Kelley JM, Gocayne JD, Dubnick M, Polymeropoulos MH, Xiao H, Merril CR, Wu A, Olde B, Moreno RF, Kerlavage AR, McCombie WR, Venter JC. Complementary DNA sequencing: expressed sequence tags and human genome project. Science 1991;252:1651–1656.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol 1990;215:403–410.

Ambros V. A hierarchy of regulatory genes controls a larva-to-adult developmental switch in C. elegans. Cell 1989;57:49–57.

Anderson L, Seilhamer J. A comparison of selected mRNA and protein abundances in human liver. Electrophoresis 1997;18: 533–537.

Augenlicht LH, Wahrman MZ, Halsey H, Anderson L, Taylor J, Lipkin M. Expression of cloned sequences in biopsies of human colonic tissue and in colonic carcinoma cells induced to differentiate in vitro. Cancer Res 1987;47:6017–6021.

Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, Aach J, Ansorge W, Ball CA, Causton HC, Gaasterland T, Glenisson P, Holstege FCP, Kim IF, Markowitz V, Matese JC, Parkinson H, Robinson A, Sarkans U, Schulze-Kremer S, Stewart J, Taylor R, Vilo J, Vingron M. 2001. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. Nat Genet 2001;29: 365–371.

Cascorbi I. Genetic basis of toxic reactions to drugs and chemicals. Toxicol Lett 2006;162:16–28.

Coen M, Ruepp SU, Lindon JC, Nicholson JK, Pognan F, Lenz EM, Wilson ID. Integrated application of transcriptomics and metabonomics yields new insight into the toxicity due to paracetamol in the mouse. J Pharm Biomed Anal 2004;35: 93–105.

Cunningham MJ. Genomics and proteomics: the new millennium of drug discovery and development. J Pharmacol Toxicol Meth 2000;44:291–300.

Cunningham MJ, Dao D. Microarrays-fabricating. In: Fuchs J, Podda M, editors. *Encyclopedia of Diagnostic Genomics and Proteomics*. New York, NY: Marcel Dekker; 2005. p 819–823.

Cunningham MJ, Shah MU. Chapter 3.1 Toxicogenomics. In: Gad SC, editor, *Handbook of Pharmaceutical Biotechnology*. Hoboken, NJ: John Wiley & Sons, Inc.; 2007. p 229–251.

Fiehn O. Metabolomics—the link between genotypes and phenotypes. Plant Mol Biol 2002;48:155–171.

Gao X, LeProust E, Zhang H, Srivannavit O, Gulari E, Yu P, Nishiguchi C, Xiang Q, Zhou X. A flexible light-directed DNA chip synthesis gated by deprotection using solution photogenerated acids. Nucleic Acids Res 2001;29:4744–4750.

Gygi SP, Rochon Y, Franza BR, Aebersold R. Correlation between protein and mRNA abundance in yeast. Mol Cell Biol 1999;19: 1720–1730.

Ingelman-Sundberg M. Genetic polymorphisms of cytochrome P450 2D6 (CYP2D6): clinical consequences, evolutionary aspects and functional diversity. Pharmacogenomics J 2005;5: 6–13.

Kalow W. Familial incidence of low pseudocholinesterase level. Lancet 1956;2:576–577.

Kane MD, Jatkoe TA, Stumpf CR, Lu J, Thomas JD, Madore SJ. Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays. Nucleic Acids Res 2000;28:4552–4557.

Kleno TG, Kiehr B, Baunsgaard D, Sidelmann UG. Combination of 'omics' data to investigate the mechanism(s) of hydrazine-induced hepatotoxicity in rats and to identify potential biomarkers. Biomarkers 2004;9:116–138.

Kulle B, Schirmer M, Toller MR, Suk A, Becker C, Tzvetkov MV, Brockmoller J, Bickeboller H, Hasenfuss G, Nurnberg P, Wojnowski L. Application of genomewide SNP arrays for detection of simulated susceptibility loci. Hum Mutat 2005;25: 557–565.

Lander ES, and over 50 other authors; International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. Nature 2001;409:860–921.

Lee RC, Feinbaum RL, Ambros V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 1993;75:843–854.

Lema C, Cunningham MJ. MicroRNAs and their implications in toxicological research. Toxicol Lett 2010;198:100–105.

Lockhart DJ, Dong H, Byrne MC, Follettie MT, Gallo MV, Chee MS, Mittmann M, Wang C, Kobayashi M, Horton H, Brown EL. Expression monitoring by hybridization to high-density oligonucleotide arrays. Nat Biotechnol 1996;14:1675–1680.

Maxam AM, Gilbert W. A new method for sequencing DNA. Proc Natl Acad Sci 1977;74:560–564.

McDonald WH, Yates JR. Shotgun proteomics and biomarker discovery. Dis Markers 2002;18:99–105.

Moka D, Vorreuther R, Schicha H, Spraul M, Humpfer E, Lipinski M, Foxall PJD, Nicholson JK, Lindon JC. Biochemical classification of kidney carcinoma biopsy samples using magic-angle-spinning ^{1}H nuclear magnetic resonance spectroscopy. J Pharm Biomed Anal 1998;17:125–132.

Robertson DG, Reily MD, Sigler RE, Wells DF, Paterson DA, Braden TK. Metabonomics: evaluation of nuclear magnetic resonance (NMR) and pattern recognition technology for rapid *in vivo* screening of liver and kidney toxicants. Toxicol Sci 2000;57:326–337.

Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci 1977; 74:5463–5467.

Schnackenberg LK, Beger RD. Monitoring the health to disease continuum with global metabolic profiling and systems biology. Pharmacogenomics 2006;7:1077–1086.

Schnackenberg L, Jones RC, Thyparambil S, Taylor JT, Han T, Tong W, Hansen DK, Fuscoe JC, Edmondson RD, Beger RD, Dragan YP. An integrated study of acute effects of valproic acid in the liver using metabonomics, proteomics, and transcriptomics platforms. OMICS 2006;10:1–14.

Shastry BS. 2005. Genetic diversity and new therapeutic concepts. J Hum Genet 2005;50:321–328.

Singh-Gasson S, Green RD, Yue Y, Nelson C, Blattner F, Sussman MR, Cerrina F. Maskless fabrication of light-directed oligonucleotide microarrays using a digital micromirror array. Nature Biotechnol 1999;17:974–978.

Southern E. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 1975;98: 503–517.

Thomas PS. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc Natl Acad Sci 1980;77: 5201–5205.

Towbin HH, Staehelin T, Gordon J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc Natl Acad Sci 1979;76:4350–4354.

Wightman B, Ha I, Ruvkun G. Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. Cell 1993;75:855–862.

Yang W, Lee DY, Ben-David, Y. The roles of MicroRNAs in tumorigenesis and angiogenesis. Int J Physiol Pathophysiol Pharmacol 2011;3:140–155.

5

TOXICITY OF THE HEMATOPOIETIC SYSTEM

SHERILYN A. GROSS

This chapter describes toxicities affecting blood. The following subjects are covered:

- Origins of blood cell types
- Assays used for analyzing hematoxicity
- Diseases of the hematopoietic system
- Agents known to be hematoxic

The hematopoietic system is the most dynamic organ system in the human body. All hematopoietic or blood cell lineage develop within the bone marrow microenvironment and each cell type is groomed toward maturation within the niches of the bone marrow by exposure to specific growth factors and cytokines, as well as cell-to-cell and cell-to-stroma interactions. Once the hematopoietic cell reaches a specific stage of maturation, the marrow releases it into the peripheral blood where it further matures into a terminally differentiated cell that performs a unique cell function. For example, one cell function for a mature red blood cell (RBC) is to carry oxygen from the lung through the peripheral blood and deliver it to the tissues. A cell function for a platelet is to release proteins necessary for clot formation and a mature neutrophil's primary function is travel to inflamed tissues where they protect the body against bacterial infection.

Despite the common origin of hematopoietic cells, there is a quantitative and qualitative equilibrium between all blood cell lineages under normal physiological conditions. If the cell does, indeed, reach terminal differentiation, each cell lineage has a normal quantitative range of mature cells in the peripheral blood and a finite lifespan. For example, RBCs normally range from 4.2 to 5.9×10^6 cells/μl with an average lifetime of 120 days. The normal range for platelets in peripheral blood is $140–440 \times 10^3$ cells/μl and a lifespan of 14 days, whereas neutrophils have a normal range of 35–85 cells/μl in peripheral blood with a lifespan of only 12 h. Therefore, the bone marrow is continually producing all blood cell lineages in varying quantities in order to maintain physiological homeostasis.

The continuous proliferation and development of all blood cell lineages within the niches of the bone marrow microenvironment makes the bone marrow susceptible to toxins, especially toxins that target proliferating cells. Toxic influences such as chemical agents, irradiation, or severe infection usually lead to a greater or lesser degree of suppression of all blood cell lineages. The most extreme result of a toxic effect is the suppression of all hematopoietic cell series. An examination of hematopoiesis will aid in understanding the necessity of continuous proliferation, and, subsequently, the vulnerability of the hematopoietic system to toxic effects.

5.1 HEMATOPOIESIS: ORIGINS OF HEMATOPOIETIC CELLS

Hematopoiesis is the process by which mature blood cells of distinct lineages (e.g., platelets, erythrocytes, granulocytes, monocytes, and lymphocytes) are produced from pluripotent hematopoietic stem cells (HSCs). It is a dynamic process with respect to both lineage decisions and origin during development. During fetal development, hematopoietic activity sequentially progresses from the yolk sac to the liver, from the liver to the spleen, and from the spleen to the bone marrow. By age 20, most of the long bones have lost their hematopoietic ability; thus continuous cell development

Principles of Toxicology: Environmental and Industrial Applications, Third Edition. Edited by Stephen M. Roberts, Robert C. James and Phillip L. Williams.

in the adult occurs in the marrow of the vertebrae, sternum, ribs, and pelvis.

The bone marrow fills a three-dimensional architecture created by the cancellous tissue normally present in the interior of the bone. The trabecular pattern is an irregular meshwork of stress-related struts within the cancellous bone. The bone marrow vasculature is derived from bone artery, which branches into a capillary network of thin-walled sinusoids. The sinusoids are lined with endothelial cells supported by a discontinuous layer of reticular cells that synthesize reticulin fibers, which form an addition layer of meshwork to support hematopoietic cell development. The discontinuous layer of reticular cells is thought to provide a route for mature hematopoietic cells to enter the bloodstream. The extracellular matrix contains collagen, laminin, and fibronectin, which help to facilitate adhesion of hematopoietic cells to bone marrow stroma. If you can envision the architecture within the bone marrow, then it is easy to imagine this three-dimensional space as a perfect home for the formation of individual niches devoted to the development of specific hematopoietic cell lineage. For additional reading on niche regulation of bone marrow cells, see Renstrom et al. (2010).

The stimulus for understanding and characterizing HSC was the clinical need for cells capable of protecting humans exposed to lethal doses of irradiation. Death due to radiation poisoning was associated with bleeding and infections; however, the mechanism by which these symptoms occurred was initially unknown. The first insight came with the observation that lead shielding of hematopoietic tissue prevented death from lethal doses of radiation. Following this observation, scientists noted that intravenous infusions of syngeneic marrow after radiation also prevented death by repopulating the bone marrow with all hematopoietic cell lineages. In 1961, the first assay for HSC, known as the colony-forming unit–spleen (CFU-S) assay was developed so that scientists could examine the functional characteristics of HSC. Using this assay, investigators observed that primitive bone marrow progenitors could give rise to hematopoietic colonies in the spleen of lethally irradiated mice. Through a series of experiments, the results suggested that bone marrow may contain highly proliferative progenitor cells capable of forming colonies of myeloid, erythroid, and megakatyocytic cells. Thus scientists proposed that the HSC must exist within the bone marrow and that these cells are capable of multilineage differentiation as well as self-renewal. These findings lead to the current view that in order for HSC to sustain hematopoiesis throughout an individual's lifetime, HSC must be capable of (i) maintenance in a noncycling state, (ii) self-renewal to generate additional HSCs, and (iii) production of progenitor cells with more limited developmental potential.

Due to the fact that the process of self-renewal and differentiation are tightly coupled within the progenitor cell pool, several models have been proposed for hematopoiesis. The models for self-renewal consider various scenarios. First, there may be multiple lineages of HSC with intrinsic differences in self-renewal and differentiation in response to the microenvironment. Second, the stem cell pool may be organized into at least three distinct populations with discrete self-renewal potential. Finally, the stem cell pool may represent a continuum of HSC with the potential for self-renewal. In addition to self-renewal, inductive and stochastic models have been proposed to interpret hematopoietic commitment and differentiation. The inductive model proposes that the binding of hematopoietic growth factors to their receptors induce differentiation, and that lineage choice is determined by multipotent hematopoietic progenitor cells (HPC) by activating the transcription of lineage-specific gene programs. In contrast, the stochastic model proposes that multipotent HPC do not require exposure to external stimuli to undergo lineage commitment. Instead, this model suggests that hematopoietic growth factors simply permit the proliferation of intrinsically committed cells and the subsequent expression of mature phenotype. It is possible that both the inductive and the stochastic models represent HPC commitment and differentiation. Regardless of the model, it is well established that growth factors and growth factor receptor expression play a key role in hematopoietic lineage commitment, growth, and differentiation.

Myelopoiesis

The typical language used to describe terminally differentiated cells in the peripheral blood includes RBCs, platelets, and white blood cells (WBCs). However, the differentiation process and terminology associated with blood cell maturation is more complicated. Blood cell maturation is referred to as poiesis—"the formation or production of."

Myelopoiesis is the production of granulocytes, erythrocytes, macrophages, and megakaryocytes from the multipotent granulocyte erythrocyte–macrophage megakaryocyte-colony-forming unit (GEMM-CFU).

Erythropoiesis

Erythropoiesis refers to the formation of RBCs (also called erythrocytes). The hormone erythropoietin, produced in the kidneys, controls the process of erythropoiesis. Erythropoiesis occurs within the niches of the bone marrow architecture and, as discussed previously, originates from the multipotent GEMM stem cell. Further commitment to the erythroid lineage includes the blast-forming unit–erythrocyte (BFU-E), which then matures into the colony-forming unit–erythrocyte (CFU-E). The proerythroblast is the earliest distinguishable red cell precursor in the bone marrow and is characterized by a dense nucleus that turns deep blue with Romanowsky staining. The daughter cell is termed a basophilic erythroblast, which has a smaller nucleus than the parent cell. These

erythroid precursors only occur in the peripheral blood under pathological conditions. Mitosis generates more mature red blood precursor cells called polychromatic erythroblasts characterized by grayish blue cytoplasm and the potential to divide. Orthochromatic erythroblasts are distinguished from the parent cell by cytoplasm with a pink hue (i.e., contains hemoglobin (Hgb)) and are no longer able to divide. The nuclei of the orthochromatic erythroblast gradually condense into small black spheres without structural definition that are eventually expelled from the cell. The now enucleated young erythrocytes contain an abundant amount of ribosomes that precipitate into reticular ("net-like") structures, hence the name reticulocytes. Reticulocytes are still considered immature erythrocytes and are only in the peripheral blood for approximately 1 day before they transform into mature erythrocytes. An increased number of reticulocytes in the blood is known as recticulocytosis and is indicative of a condition that has stimulated erythropoiesis such as anemia or blood loss. As previously stated, erythrocytes have a life span of approximately 120 days before their membranes weaken and the reticuloendothelial phagocytic cells in the spleen remove them from the blood.

Mature erythrocytes have a discoid shape and are loaded with the heme-containing protein, Hgb, which is responsible for their unique function. Erythrocytes' function is to transport oxygen to tissues in the arterial vasculature and carry carbon dioxide back to the lungs for elimination in venous vasculature. Erythrocytes also provide a blood pH buffering capacity by converting carbon dioxide to carbonic acid via the enzyme carbonic anhydrase. Erythrocytes possess an outer membrane that is supported by a complex cytoskeletal system essential for squeezing through the smaller capillaries of the vascular system. Not surprisingly, if this cytoskeleton is damaged in response to reactive oxygen species, chemical attack, or hereditary disease, the changes may lead to membrane fragility and hemolysis.

Thrombopoiesis

Thrombopoiesis is the technical term for the development of thrombocytes (platelets). The hormone thrombopoietin, produced by the liver, bone marrow stromal cells, and various other organs, is the primary regulator of platelet production. Platelets are derived from the multipotent GEMM stem cell and thrombopoiesis proceeds independently toward the development of megakaryoblast cells when the granulocyte–macrophage stage of maturation is complete. The megakaryoblast reduplicates its nuclear and cytoplasm components up to seven times without cell division, and each reduplication causes an increase in the number of complete chromosome sets, nuclear lobulation, and cell size. Thus the megakaryocyte is an extremely large, lineage-committed precursor cell that is easily recognized by its size and multilobed nucleus within the niches of the bone marrow. Platelets are formed from the cytoplasmic fragmentation of the megakaryocyte within the bone marrow and are identified in the peripheral blood as small disc-shaped anucleated cells.

The normal number of platelets in the peripheral blood ranges from 140 to 440×10^3 cells/μl with a lifespan of 14 days. Mature platelets have a well-developed cytoskeleton with bands of microtubules. In addition, the mature cytoplasm has granules that contain coagulation factors, serotonin, acid hydrolases, and peroxidase. These components aid in the platelets' ability to serve as the first line of defense against damage to blood vessels. Once the epithelial lining of a blood vessel is compromised, platelets adhere to collagen by interacting with coagulation factor. Platelet actin, myosin, and microtubules cause reversible platelet adhesion along a broad surface of the damaged endothelium. Platelets release the contents of their granules and begin the synthesis of thromboxane. Combining with thromboxane, ADP and Ca^{2+} ions mediate the adhesion of other platelets. Platelet phospholipids activate the blood clotting cascade, leading to the formation of fibrin.

Granulopoiesis

Neutrophils, eosinophils, and basophils are all considered as granulocytes because at maturation the cytoplasm of these cell types contains prominent granules. All granulocytes are derived from a common myeloid-committed progenitor cell and the first recognizable granulocytic precursor cell is the myeloblast.

Neutrophil Maturation Myeloblast maturation into a neutrophil takes approximately 7–8 days. The differentiation process includes the subsequent stages of promyelocyte, myelocyte, metamyelocyte, band cell, and finally, the segmented neutrophil. The band cell has a horseshoe-shaped nucleus whereas the segmented neutrophil can be identified by its multilobed nucleus and abundant cytoplasm. Although no further division takes place, cellular acquisition of chemotactic ability, receptor expression, and complement are features of neutrophil maturation. Structurally mature neutrophils remain in the bone marrow for an additional 5 days prior to their release into the bloodstream. Neutrophils only circulate in the peripheral blood for approximately 6 h before they migrate into tissues where they can survive for an additional 2–5 days. Interestingly, some neutrophils are stored in the bone marrow and are capable of quick mobilization in response to bacterial infection.

Neutrophils have a major role in the phagocytosis of bacteria and dead cells; therefore, they contain acid hydrolase, myeloperoxidase, alkaline phosphatase, and antibacterial and digestive agents, among many other substances. Neutrophils are attracted to areas of infection or tissue damage by the release of degrading cellular chemicals (i.e., chemotaxins). They migrate through the endothelium and, once in the tissue,

the neutrophils move by way of actin filaments. Neutrophil phagocytosis expends all the neutrophils' cellular energy, and they die soon after. The collection of dead neutrophils, tissue fluid, and degradation material is termed *pus*.

Basophil Maturation Basophils share a common bone marrow precursor cell with the other granulocytic cells up to the myeloblast stage. Basophilic differentiation is analogous to neutrophil maturation. Mature basophils are characterized by large, basophilic (blue-stained) granules and are considered precursors to tissue mast cells. Mast cells migrate to supportive tissue, like below the epithelia, around blood vessels, and in the lining of serous cavities. They are long-lived and can proliferate in tissue in response to T lymphoctyes.

The granules of both basophils and mast cells contain sulfated protoglycans, heparin, chondroitin sulfate, histamine and leukotriene 3. Both basophils and mast cells have surface receptors that are highly specific for allergens. Exposure to allergens results in rapid exocytosis of their granules and release of histamine and vasoactive mediators resulting in an immediate hypersensitivity reaction.

Eosinophil Maturation Eosinophils also share a common progenitor cell with other granulocytes and their differentiation pathways diverge after the myeloblast stage as well. Eosinophils are distinguishable from neutrophils at the myelocyte stage by the appearance of eosinophilic (i.e., red-stained) granules. Eosinophils remain in the bone marrow for several days and then circulate in the peripheral blood for 3–8 h before migrating into the skin, lungs, and gastrointestinal tract. Interestingly, eosinophils have a marked diurnal pattern with maximum numbers circulating in the peripheral blood in the morning and minimal numbers in the afternoon.

Eosinophilic granules contain hydrolytic lysosomal enzymes and peroxidase, acid phosphatase, large concentrations of aryl sulfatase, and an alkaline protein referred to as *major basic protein*. Perhaps one of their main functions of eosinophils is that they increase in number in response to parasitic infection. They can also phagocytize small molecules or undergo degranulation in response to bacterial products and complement components. In addition, eosinophils are attracted to products of mast cells and respond to the localized destruction of mast cell granules by neutralizing histamines and releasing prostaglandins thought to inhibit mast cell degranulation.

Monopoiesis

Monocytes are derived from the multipotent GEMM stem cell in the bone marrow, and differentiation proceeds independently toward the development of macrophage when the granulocyte–macrophage stage of maturation is complete. The monoblast and the promonocyte are two morphological distinguishable precursors. At least three cell divisions occur before the monocyte maturation is reached. Mature monocytes leave the bone marrow soon after maturation and circulate in the peripheral blood for approximately 3 days before migrating to the tissues. Circulating monocytes are large, motile, phagocytic cells, and contain numerous lysosomal granules and cytoplasmic vacuoles. The granules contain acid phosphatase, aryl sulfatase, and peroxidase analogous to the components found in neutrophilic granules. Monocytes have numerous small pseudopodia extending from their cellular membrane and respond chemotactically to necrotic material, invading microorganisms, and inflammation. When monocytes leave the bloodstream, they are referred to as *macrophages* and are unable to reenter the circulation.

The monocyte–macrophage system of circulating monocytes and tissue macrophages are found in both the free and the fixed forms. This cell system also includes the Kupffer cells of the liver, pulmonary alveolar macrophages, and dendritic antigen presenting cells.

Myelopoietic Growth Factors Multiple growth factors contribute to cellular maturation of each lineage at various stages of cell differentiation. These growth factors were defined historically, by trial and observation of the type of lineage-committed HPC that were stimulated to produce colonies. A great deal of information has been accumulated regarding chromosome mapping, protein structure, cellular origin, and cellular functions of these growth factors, as well as other interleukin cytokines and their respective receptor expression—all of which contribute to lineage commitment, cellular differentiation, and cell function. A detailed discussion of these growth factors, cytokines, and receptor expression are beyond the scope of this book chapter. For additional information, please refer to *Jandl's* (1996) Textbook of Hematology.

Lymphopoiesis

The bone marrow is also the site for production of both B and T cell lymphocytes. A common myeloid and lymphoid stem cell was demonstrated in irradiated mice transplanted with HSCs and evidenced by the repopulation of the irradiated bone marrow with all hematopoietic cell lineages. Although both the myeloid and the lymphoid cell lineages are thought to originate from a pluripotent HSC, commitment toward the lymphoid lineage is thought to occur in a pluripotent lymphopoietic bone marrow cell.

T lymphocytes migrate from the fetal bone marrow to the thymus where they undergo three to four additional divisions before their release into the peripheral blood and lymph nodes. Mature T lymphocytes are responsible for cell-mediated immune response. In contrast to T lymphocytes, B lymphocyte progenitor and precursor cells reside and differentiate in the bone marrow. Most B cells leave the

bone marrow before full maturation and travel to the spleen and lymph nodes where they acquire specific receptors and surface immunoglobulins such as IgM and IgD. Adult bone marrow retains very few B lymphocytes. Mature B lymphocytes are responsible for humoral immunity mediated by the secretion of antibodies.

Lymphocytes are the most numerous type of white cell in the peripheral blood. Most lymphocytes are small with only 3% considered large activated lymphocytes on route to target tissues. The lymphocyte has a kidney-shaped nucleus with dense chromatin. It is important to note that small lymphocytes in the peripheral blood cannot be distinguished as B or T lymphocytes by morphology or staining and are not the functional end form, but undergo transformation in response to specific immunological stimuli. However, plasma cells, which are a differentiated form of B lymphocytes, can be visualized in the peripheral blood by the morphological trademark "pin–wheel" chromatin pattern, which is reflective of active immunoglobulin synthesis.

Toxicity associated with B and T lymphocytes is technically termed *immunotoxicity* and is usually discussed independent of hematoxicity. Lymphopoiesis is discussed here only for the purpose of inclusion of lymphocytes as a component of WBCs in the peripheral blood. Immunotoxicity will not be discussed in this chapter.

5.2 HEMATOTOXICITY: ASSAYS USED FOR ANALYSIS

There is a quantitative and qualitative equilibrium between all blood cells under normal conditions. This equilibrium is regulated by a variety of growth factors, cytokines, and cellular communication in order to ensure physiological balance between blood cell production and degradation. There are a number of assays that are routinely used to determine the status of blood cell equilibrium. For example, a complete blood count (CBC) and a blood chemistry profile are common automated assays performed with peripheral blood as part of a wellness examination in order to establish baseline parameters for a multitude of body processes (e.g., red and white cell counts, clotting factors, liver enzymes, serum lipids, electrolyte levels, and metabolic by-products). In turn, disturbances in bone marrow function are accompanied by changes in the composition of peripheral blood cells and blood proteins. CBC and blood chemistry studies are also used for clinical diagnosis when disorders of the blood and bone marrow are suspected.

In addition to routine hematology testing, the World Health Organization introduced the *Classification of Tumours of the Haematopoietic and Lymphoid Tissues* in 2001 (Jaffe, 2001). The objective was to provide pathologists, oncologists, and geneticists with unified criteria for diagnosis of human neoplasms defined according to a combination of composition of the blood and bone marrow, cellular morphology, immunophenotype, genetic features, and clinical features. This established criterion was based on the state of the science at that time and the use of state-of-the-art diagnostic tools. Since 2001, the scientific knowledge of cell biology and disease has evolved and the resolution of diagnostic tools has improved. As a result, in 2008, WHO published a second edition of the classification system for tumors of the hematopoietic and lymphoid tissue (Swerdlow, 2008).The complete references for Jaffe (2001) and Swerdlow (2008) are found in the reference section of this chapter.

Complete Blood Count

A CBC is an automated measurement of the various cell lineages circulating in the peripheral blood at a given moment. The automated cell count has largely replaced the "manual" cell count as an inexpensive, easy, and rapid way to obtain information about the hematological system as well as other organ systems. Normal CBC values are affected by several factors, including age, gender, metabolic activity, circadian rhythms, and nutritional status, as well as blood sampling technique, storage, and counting method. For this reason, laboratories provide a normal range in which 95% of the values are found in a clinically normal group of individuals. Overlap between normal and pathological data is expected to some degree. Therefore, values that fall within the borderline areas must be interpreted with respect to the reference range of normal individuals that closely resemble the patient. For example, the reference range for leukocytes varies between newborns, toddlers, young children, and adults. Once the appropriate reference range is established, the CBC provides the current status of erythrocytes, WBCs (leukocytes), and platelets.

Peripheral Blood Smear

The CBC is usually accompanied by a differential blood cell count performed with a peripheral blood smear. The differential measures the percentage of each type of leukocyte present in the same specimen. The peripheral blood smear is prepared manually by spreading one drop of blood sample across a glass slide to achieve a thin smear. The blood is smeared from one end of the glass slide to the other, ending in a "feathered" edge. Dried peripheral blood smears are stained with a mixture of acidic and basic stains so individual WBCs can be distinguished from each other by their cellular components. The cells are counted manually under a microscope.

Table 5.1 contains an example of a CBC and accompanying differential. The significance associated with each test is discussed separately.

TABLE 5.1 Normal Range and Mean Values For CBC in Adults

Cell Type	Mean Value	Normal Range
Leukocytes (10^6/l)[a]	7000	4,300–10,000
Neutrophils bands (%)	2	0–5
Neutrophil segs (%)	60	35–80
Eosinophils (%)	2	0–4
Basophils (%)	0.5	0–1
Monocytes (%)	4	2–6
Lymphocytes (%)	30	20–50
Erythrocytes (10^{12}/l)[a]	Male: 5.4	4.6–5.9
	Female: 4.8	4.2–5.4
MCH (Hb_E pg)	29	26–32
MCV (fl)[a]	87	77–99
MCHC (g/dl)	33	33–36
RDW (μm)	7.5	
Reticulocytes (%)	Male: 16	8–25
	Female: 24	8–40
Hgb (10 g/dl)	Male: 15	14–18
	Female: 13	12–16
Hct (%)	Male: 45	42–48
	Female: 42	38–43
Platelets (10^3/μl)	180	140–440

Hct, hematocrit; Hgb, hemoglobin; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; MCV, mean corpuscular volume; RDW, red blood cell distribution width.
Reference range: 95% of the population.
[a]SI units give the measurements per liter.

Leukocytes (White Blood Cell Count) The leukocyte count refers to the total number of WBCs in 1 mm^3 of peripheral blood. As with the CBC in general, the normal range for leukocytes varies with age. The normal range for leukocytes in a newborn is 9,000–30,000 × 10^6/l and in toddlers the range is from 6,200 to 17,000 × 10^6/l. Although there is a wide range for total number of leukocytes, there are many causes for abnormal values. In an adult, the critical values may be below 2,500 10^6/l (leukopenia) and above 10,000 10^6/l (leukocytosis). An increase in total WBC usually indicates infection, inflammation, necrotic tissue, or leukemia neoplasm, although emotional stress or physical trauma can also cause an increase in WBC. A decrease in total number of leukocytes is most often associated with bone marrow failure induced by antineoplastic agents, radiation, diseases that infiltrate the bone marrow, overwhelming infections, nutritional deficiencies, and autoimmune diseases.

The differential WBC count refers to the percentage of each subtype of leukocyte within the total number of leukocytes in a given sample. Leukocytes range from 8 to 20 μm in size. There are five types of leukocytes that compose the WBC differentiation, and these cell subtypes are characterized by the presence or absence of granules. The agranulocytes are lymphocytes and monocytes. The granulocytes are the neutrophils, eosinophils, and basophils. An increase in the percentage of one cell type leads to a decrease in the percentage of another cell type. Neutrophils and lymphocytes make up 75–90% of the total number of leukocytes. When neutrophil production is stimulated, immature forms of neutrophils (i.e., bands) enter the circulation resulting in less mature forms of neutrophils. In the clinical setting, this is referred to as a *shift to the left.*

Eosinophils Increased production of eosinophils is associated with allergic reaction, as these cells are capable of phagocytosis of antigen–antibody complexes. When the allergic reaction subsides, the eosinophil count decreases. Increase in eosinophils is also seen with drug-induced hypersensitivity. These drugs include various antibiotics, gold preparations, hydantoin derivatives, phenothiazines, and dextrans. Hypereosinophilia syndrome, with eosinophil values greater than 40% of the total leukocyte count, is seen with splenomegaly, heart defects, and pulmonary infiltrates. When hypereosinophilia presents clinically in combination with these diseases, it is classified as either an autoimmune disease or myeloproliferative disorders, or somewhere in between. Increase in eosinophil production is also associated with parasitic infection. Acute eosinophilic leukemia is rare.

Basophils Increase in basophil production (<2–3%) is rare, although it can be associated with allergic reactions to food, drugs, or parasites and is usually seen in clinical cases when eosinophils are also involved. Basophilia is seen with infectious diseases such as tuberculosis and chickenpox and with metabolic diseases such as hyperlipidemia. Proliferation of basophils is associated with myeloproliferative neoplasms, although acute basophilic leukemia is rare.

Monocytes Monocytes function as phagocytic cells inside and outside of the circulatory system. An increase in the monocytic cell population above 7% in the differential blood cell count is indicative of an immune defense reaction. Monocytosis occurs in the case of infection although it usually presents near the end of the infection. Chronic monocytosis is seen especially in brucellosis and tuberculosis. Monocytosis can also be part of a noninfectious response such as Crohn's disease and ulcerative colitis. Interestingly, monocytosis often occurs in response to a disseminating neoplasm such as carcinoma of the bronchi and breast. Perhaps not surprisingly, monocytosis is prominent in myeloproliferative disorders such as chronic myelomonocytic leukemia (CMML) and acute monocytic leukemia. In fact, monocytic leukemia is associated with a sharp rise in peripheral blood monocyte counts accompanied by a drop in absolute counts in other cell types noted in the differential.

Lymphocytes Alterations in lymphocyte counts are most often associated with viral infections or in diseases of the lymphatic system. A spontaneous decrease in lymphocyte count is seen only in some rare congenital diseases such as agammaglobulinemia and chromosome 22q11 deletion

syndrome. Further, low lymphocyte counts are seen in some systemic diseases such as Hodgkin's disease and active AIDS. As previously stated, an increase in the percentage of one cell type leads to a decrease in the percentage of another cell type. An example of this is a toxic response in the neutrophil cell series that manifests as a benign increase in the absolute number of lymphocytes. The technical term for this abnormality is *agranulocytosis*, which refers to an increase in WBCs without granules. There are several different classifications for neutropenia and agranulocytosis, such as drug hypersensitivity reaction, infection, autoimmune neutropenia, congenital or familial neutropenia, and neutropenia secondary to a bone marrow disease.

Hemoglobin and Hematocrit The Hgb concentration is the measure of the total amount of Hgb in the peripheral blood and is a reflection of the number of erythrocytes. The hematocrit (Hct) is a measure of the percentage of total blood volume that is made up of RBCs and is a close reflection of both the Hgb and the RBC value. The Hct value is usually three times the Hgb concentration. A decrease in both Hgb and Hct values are associated with anemia; an increase is associated with erythrocytosis.

Erythrocytes (Red Blood Cell Count) The erythrocyte count represents the number of RBCs in 1 mm^3 of peripheral blood. The normal range for erythrocytes tends to decrease with increasing age. The normal range for erythrocytes in a newborn spans from 4.8 to 7.1×10^{12}/l and in young children from 4.0 to 5.5×10^{12}/l. An adult female has normal RBC values ranging from 4.2 to 5.4 whereas an adult male ranges from 4.7 to 6.1×10^{12}/l. Other causes of variations in RBC counts include a decreased count during pregnancy, an increased count at high altitudes, and an increased or decreased count depending upon hydration status.

Intravascular abnormalities can shorten the lifespan of RBCs by causing trauma to the blood cell membrane. These abnormalities include artificial heart valves and peripheral vascular atherosclerosis. An enlarged spleen can also cause an early, inappropriate destruction of erythrocytes. When the RBC values decrease by 10% of the expected normal value, the individual is labeled anemic. Additional causes of low RBC values are hemorrhage, hemolysis, hemaglobinopathy, advanced cancer, bone marrow fibrosis, chemotherapy, renal failure, multiple myeloma, leukemia, and dietary deficiencies. Abnormal increases in RBC values are related to congenital heart disease, polycythemia vera (PV), and pulmonary fibrosis.

Red Blood Cell Indices RBC indices provide information about RBC size (MCV and RDW), weight (mean corpuscular hemoglobin (MCH)), Hgb concentration (mean corpuscular hemoglobin concentration (MCHC)), and maturation (reticulocyte count). RBCs are approximately 8 mm in size, which allows them to pass through the tiny capillaries. Cell size is indicated by terminology such as normocytic, microcytic, and macrocytic. In turn, Hgb content is indicated by the terms normochromic, hypochromic, and hyperchromic. Extremely elevated WBCs and abnormal number of immature RBCs circulating in the peripheral blood can actually affect the RBC indices.

Mean Corpuscular Volume Mean corpuscular volume (MCV) is a measure of the average size of a single RBC and is used in classifying anemias. An increase in MCV is associated with abnormally large RBCs typically seen with liver disease, antimetabolite therapy, alcoholism, pernicious anemia, and folic acid deficiency. A decrease in MCV is associated with abnormally small RBC and is seen with iron deficiency anemia and thalassemia.

Red Blood Cell Distribution Width Red blood cell distribution width (RDW) value is indicative of the variation in RBC size. The RDW is important in determining the degree of variability and abnormality in RBC size (e.g., anisocytosis). An increase in RDW is associated with iron deficiency anemia, B_{12} or folate deficiency anemia, hemolytic anemia, and posthemorrhagic anemias.

Mean Corpuscular Hemoglobin MCH is a measure of the average amount of Hgb within an RBC. Macrocytic cells have more Hgb content whereas microcytic cells have less Hgb. As such, an increase in MCH is associated with macrocytic and hyperchromic anemia whereas a decrease in MCH is associated with microcytic and hypochromic anemia.

Mean Corpuscular Hemoglobin Concentration MCHC is a measure of the average concentration percentage. This value is used to determine whether the RBCs have increased, decreased, or whether there are normal concentrations of Hgb (e.g., hypochromic, hyperchromic, normochromic). Increases in MCHC are associated with spherocytosis, intravascular hemolysis, and cold agglutins. A decrease in MCHC is associated with iron deficiency anemia and thalassemia.

Reticulocyte Count A reticulocyte is an immature RBC and the reticulocyte count can be used to determine bone marrow function and erythropoietic activity. Normally, there are a few reticulocytes circulating in the peripheral blood. An increase in reticulocytes is indicative of the bone marrow producing an increased number of RBCs, which usually occurs in response to anemia. However, if Hgb is normal, then an increase in the number of reticulocytes in the peripheral blood can be associated with increased production of RBCs in response to ongoing loss of red cells from hemolysis or hemorrhage. A normal or low reticulocyte count in an anemic patient indicates that the bone marrow's response to the anemic condition is an inadequate

production of RBCs or the inadequate production of RBCs is causing the anemia (e.g., aplastic anemia, iron deficiency, Vitamin B_{12} deficiency).

Thrombocytes (Platelets Count) Platelet count is the number of platelets per 1 mm^3 of peripheral blood. Platelet counts of 150,000–400,000 mm^3 are considered within normal range. Counts less than 100,000 mm^3 indicate thrombocytopenia, whereas counts that are greater than 400,000 mm^3 are referred to as thrombocytosis. Hemorrhage may occur with severe thrombocytopenia especially when platelet counts drop below 20,000 mm^3. Thrombocytopenias develop secondary to bone marrow failure or from tumor infiltration of the bone marrow. They can also arise from the sequestration of platelets by the spleen or consumption of platelets due to coagulation within the vasculature. In addition, thrombocytopenias are caused by the accelerated destruction of platelets secondary to antibodies, drugs, or infections. Thrombocytopenia due to reduced platelet production is also associated with chronic alcoholism, cytostatic drugs, virus infection, and vitamin deficiency. Spontaneous thrombocytosis is often associated with malignancies such as leukemias, lymphoma solid tumors, PV, rheumatoid arthritis, and iron deficiency anemia. A platelet count greater than 1,000,000 mm^3 is referred to as *thrombocythemia.*

Platelet counts can vary with high altitude and strenuous exercise. Oral contraceptives can increase platelet count, whereas menstruation and some over-the-counter drugs such as acetaminophen, aspirin, and cimetidine can decrease platelet count. Individuals with a low platelet count may have bruising, petechiae (i.e., tiny purple or red spots on the skin resulting from small hemorrhages), nose bleeds, and bleeding of the gums.

Mean Platelet Volume The mean volume varies with total platelet production. For example, normal bone marrow will react to thrombocytopenia associated with sequestering of platelets by the spleen by releasing larger, immature platelet cells in an effort to maintain a normal platelet count. Therefore, the mean platelet volume will increase. In contrast, if the production of platelets in the bone marrow is inadequate, the released platelets are small and pyknotic (old).

For further information on peripheral blood parameters see *Wintrobe's Clinical Hematology*, *Mossby's Diagnostic and Laboratory Test References*, and *Mosby's Medical Dictionary.*

Bone Marrow Aspirate and Biopsy

The bone marrow cellularity, myeloid:erythroid (M:E) cell ratio, and a bone marrow differential cell count are simple diagnostic tools used to determine the extent of bone marrow involvement in hematotoxicity.

Bone Marrow Collection Bone marrow aspirate and biopsy are usually obtained from the upper part of the posterior iliac crest (i.e., back of the hip bone). For the core biopsy, a deep local anesthesia is used to numb the area and then a small incision is made on the skin. The core biopsy is collected using a Jamshidi, which is a combination of an obturator, a sharp hollow needle, and a hollow histology cylinder that is at least 1.5 cm in length. The instrument is twisted back and forth through the soft tissue and approximately 30 mm into the bone marrow space. The needle is removed with a twisting motion and then the compact core biopsy is extruded with a plunger. The procedure preserves the bone marrow architecture.

The bone marrow aspirate can be collected independently using a cytology needle that is guided by a stylus, which is slowly pushed through the through the soft tissue and the compact of the bone into the marrow space. Removal of the stylus is followed by withdrawal of the plunger, filling the syringe with bone marrow cells in suspension.

Bone Marrow Cellularity The distribution of cell types within the bone marrow is fairly constant. A semi-quantitative assessment of the bone marrow cellularity is determined using both the bone marrow aspirate and sections of the core biopsy. There are three crude categories used to describe the relative proportions of bone marrow cells, but the differences between the categories are somewhat arbitrary. For example, bone marrow cellularity is referred to as *hypocellular* (<40% cellular), *normocellular* (40–60% cellular), and *hypercellular* (>60% cellular). Significant changes in cellularity as a whole or in specific cell types help to support a diagnosis.

Myeloid:Erythroid (M:E) Ratio A normal M:E ratio is approximately 3:1 and is valid only as a crude index of relative cellularity. A change in this ratio could mean depression of erythropoiesis or enhancement of myelopoiesis. The M:E ratio is determined in combination with additional assays to help develop a total picture of the bone marrow status.

Marrow Differential Count A normal bone marrow differential is a morphological representation of all cell series in varying stages of maturation. Manual differential counts are prepared with microscope slides of bone marrow aspirate and high-powered oil emersion microscopy. An example of a normal bone marrow differential is illustrated in Table 5.2.

Advances in cell biology have greatly improved knowledge of cell growth and differentiation. Additionally, there have been advances in the identification of molecules that appear both on the cell surface and intracellularly during the normal process of differentiation and maturation. As a result, changes in the normal morphology, histochemistry,

TABLE 5.2 Normal Bone Marrow Differential

Cell Types	Range %
Myeloblasts	<5
Promyelocytes	1–8
Myelocytes	
Neutrophilic	5–15
Eosinophilic	0.5–3
Basophilic	<1
Metamyelocytes	
Neutrophilic	15–25
Eosinophilic	<1
Basophilic	<1
Mature myelocytes	
Neutrophilic	10–30
Eosinophilic	<5
Basophilic	<5
Mononuclear	
Monocytes	<5
Lymphocytes	3–20
Plasma cells	<1
Megakaryocytes	<5
Normoblasts	25–50

cytology, and cytogenetics are important diagnostic markers used to determine the extent of bone marrow involvement in hematotoxicity and/or hematopoietic disease.

Histochemistry Using routine staining techniques, cell histology allows for examination of cellular structures in paraffin-fixed bone marrow tissue samples, the most common of which is the H&E stain. In addition, certain dyes have affinity for chemical groups on cellular molecules. For example, lipid moieties are detected by lipid-soluble stains such as Sudan Black. Polysaccharides are identified by periodic acid–Schiff (PAS), which oxidizes glucose residues into aldehydes. A Schiff reagent (Fuchsin) is added, which reacts with the aldehyde to produce purple fuchsin. Enzyme activity can also be identified using an enzyme substrate, a cofactor, and visualizing agents.

Immunocytochemistry methods are base on antibodies raised with affinity for specific cellular molecules. Detection of the primary antibody is performed using a second antibody linked to an enzyme called horseradish peroxidase. A well-established antibody detection method, the horseradish peroxidase is developed as a colored reaction and visualized under light microscopy. Immunofluorescence techniques are based on the same antibody affinity, except fluorescence label antibodies are used to identify cellular molecules with fluorescence microscopy.

Flow Cytometry/Immunophenotyping

Immunophenotyping is based on fluorescently labeled antibodies with affinity for cluster designation (CD) molecules on the cell surface. Over the last two decades, numerous CD molecules have been identified as cellular protein configurations that are expressed on the cell surface and further characterized by functional activity. These surface molecules include, but are not limited to, growth factor and cytokine receptor molecules, adhesion molecules, ion channels, and lineage-specific molecules. Expression of many of these surface molecules occurs in all cell types and is responsible for normal cellular function (e.g., movement across ion channels). Some surface proteins alter expression as the cell matures (e.g., adhesion molecules, growth factor receptors), and some of these surface molecules play a specific role in mature cell activity (e.g., lineage-specific growth factor receptors). Depending on the specific cell type and the cell's stage in maturation, these surface molecules provide the phenotypic characteristic of the cell. It is important to note that the cellular phenotype remains fluid throughout lineage maturation until terminal differentiation. Immunophenotyping is the process of fluorescent labeling of these molecules on the cell surface to determine a cell's specific cell lineage in a specific stage of cellular maturation. Furthermore, noted changes in the normal cellular expression pattern are also used as a diagnostic tool to follow abnormalities in cellular proliferation and differentiation.

Many fluorochrome-conjugated monoclonal antibodies with affinity for specific CD molecules are commercially available for analysis of bone marrow cells in suspension. With the use of compatible fluorochromes, multiple conjugated antibodies can be combined to further characterize subpopulations of cells and to minimize the number of cells required for analysis. Fluorescently stained cells are analyzed by flow cytometric techniques in which an argon laser is used to excite the fluorochromes so that each emits spectra at an individual wavelength. The signals from forward and right-angle side scatter are collected in a logarithmic amplifier and then analyzed with software designed specifically for flow cytometric methods.

In addition to immunophenotyping, flow cytometric techniques identify intracellular protein expression, cell cycle progression, and apoptosis assays. Hematotoxins and diseases associated with the hematopoietic system can manifest as alterations in differentiation, disruption of the cell cycle, as well as induction and/or inhibition of apoptosis. Flow cytometric techniques are therefore useful in characterizing deviations from normal patterns of cell proliferation, cell differentiation, and cell death. For additional reading on flow cytometric techniques see Howard *Shapiro's Practical Flow Cytometry*.

Chromosome Analysis

Karyotyping The study of whole chromosomes is called *karyotyping*. In this assay, chromosomes are stained with a specific dye followed by chemically induced cell cycle arrest. Pairs of chromosomes are arranged by size and position of

centromeres. Karyotypes are arranged with the short arm of the chromosome on top and the long arm on the bottom. The short and long arms are referred to as *p* and *q*, respectively. In addition, the differently stained regions and subregions are given numerical designations from proximal to distal on the chromosome arms. The normal human karyotypes contain 22 pairs of autosomal chromosomes and 1 pair of sex chromosomes. Normal karyotype for a female contains two X chromosomes and is denoted as 46,XX whereas a male-carries both an X and a Y chromosome, 46,XY.

Chromosome abnormalities can be numerical, as in the presence of extra or missing chromosomes, or structural, as in translocations, inversions, large-scale deletions, or duplications. Regarding numerical abnormalities, the term *ploidy* refers to the number of chromosome pairs in a karyotype, and *aneuploidy* refers to abnormal number of chromosomes. *Trisomy* is the terminology used when three copies of a particular chromosome are present instead of two, and *monosomy* refers to a single chromosome copy.

Cytogenetics Chromosome banding employs different cytogenetic techniques to examine specific areas of the chromosome. Since each chromosome within the pair should have identical bands, comparing differences in banding patterns between chromosome pairs can pinpoint the area of the chromosome abnormalities. In G-banding, chromosomes are digested with trypsin and then identified with Giemsa stain, which binds to the phosphate regions of DNA and reveals both lightly and darkly stained chromosome sections. The lighter stained regions tend to be rich in guanine and cytosine (GC) base pairs whereas the dark regions are rich in adenine and thymidine (AT) base pairs. R-banding is the reverse of G-banding in that the dark regions are GC rich and the light regions are AT rich.

Fluorescence In Situ Hybridization Fluorescence *in situ* hybridization (FISH) techniques are used to map DNA locations on specific chromosomes using fluorescent-labeled probes that are specified to a DNA sequence in order to detect numerical and structural chromosome abnormalities. FISH analysis is usually performed *in vitro* using cultured cells of interest undergoing metaphase. A large number of FISH probes are commercially available with a variety of fluorochromes, allowing for simultaneous detection of multiple DNA loci. FISH probes are useful in the detection of known translocations, inversions, insertion, as well as microdeletions and chromosome breakpoints. For additional molecular genetics review read Wan and Ma (2012). For human cytogenetic nomenclature, read ISCN (2009).

Colony-Forming Units Assay

The colony-forming potential of HPC was established nearly 50 years ago. Colony-forming unit assays are an *in vitro* culture system that allows for an analysis of the colony-forming potential of normal, chemically treated, and diseased hematopoietic cells. This assay has evolved over time. Today, we know that different culture conditions are required for growth of primitive stem cells as compared to culture conditions required for relatively mature precursor cells. The late stages of hematopoiesis can be reproduced *in vitro* in a 14-day culture medium, whereas primitive HSCs require a two-step *in vitro* culture system: 5-week preincubation on feeder layer followed by 2 weeks in semisolid media.

Colony-forming unit assays represent the intermediate state of hematopoiesis between repopulating HSC and the morphologically identifiable features of lineage differentiation. Lymphoid cells do not proliferate in culture systems that mimic the hematopoietic environment. Thus CFU assays allow for the detection of the specific lineage commitment of amyeloid precursor cell in an *in vitro* setting.

In the clinical setting, CFU assays are used to evaluate deviation from normal colony-forming potential, such as an increase or decrease in overall colony growth, change in the M:E ratio, an increase in lineage-specific CFU, deficient hemoglobinization, and abnormally small quantities of hemoglobinized cells. While changes like these in an *in vitro* assay can be useful in differential diagnosis, the assay does not stand on its own as a definitive diagnostic tool for hematopoietic disease. For further reading, see Nissen-Druey (2005), *Hematopoietic Colonies in Health and Disease*, and Sornberger (2011), *Dreams and Due Diligence.*

5.3 DISEASES OF THE HEMATOPOIETIC SYSTEM

This section is not meant to be a comprehensive discussion or description of all diseases associated with the hematopoietic system. Instead, this section is limited to blood disorders commonly associated with hematotoxins. For the purposes of this book chapter, tables are used to help categorize classifications of hematopoietic disorders. For a comprehensive review of diseases of the hematopoietic system, please see Jandl (1996), Greer et al. (2008), Jaffe (2001), and Swerdlow (2008).

Anemia

Anemia, in the general sense, is a reduction in the concentration of RBCs and Hgb in the blood. Additional considerations associated with anemia are the total volume of RBCs as compared to the total volume of whole blood, that is, the ratio of the two parameters described earlier as the *hematocrit*. Other considerations include the rate of onset of anemia, the extent of reduction in Hgb concentration, and the body's compensatory response. The cardiovascular system's response to acute anemia is to increase the velocity of blood flow and to redistribute blood to organs that are

most vulnerable to hypoxia. In chronic anemia, the same adaptive physiological changes occur so that the body can establish a new steady state. Anemias of the peripheral blood are classified into two primary categories: increased destruction (Table 5.3) or impaired production of RBCs (Table 5.4).

Anemia resulting from another type of disorder is referred to as *secondary anemia*. Secondary anemia can occur in situations where there is a lack of iron necessary for erythropoiesis. This deficit could be due to an internal iron shift caused by an "iron pull" of the reticuloendothelial system during infections, toxic responses, autoimmune diseases, and tumors. Bone marrow analysis in secondary anemia shows a reduction in both erythrocytes and granulocytes with evidence of immature cells. This is in contrast to the phenomenon associated with exogenous iron deficiency anemia where the size and shape of the RBC has specific characteristics such as variation in size (i.e., anisocytosis), variation in shape (i.e., poikilocytosis), and bluish-gray staining due to diminished amounts of Hgb (i.e., polychromophilia).

Impaired production resulting from a defect at the pluripotential stem cells level manifests as a failure to maintain stem cell kinetics such as self-renewal and differentiation. It is important to note that impaired production can occur anywhere in the process of stem cell differentiation and then manifest at an even later stage of cell maturation. For example, if the defect occurs in a committed progenitor cell such as CFU-E, then the disease is referred to as *pure red cell aplasia*. A defect in these early stem cells can lead to inadequate hematopoiesis and a reduction of colony-forming cells in all hematopoietic cell lineages (erythrocytes, granulocytes, monocytes, and megakaryocytes). This is called *bone marrow failure*, *hypoplastic anemia*, or *aplastic anemia*. The peripheral blood parameters reflect the overall decrease in lineage production referred to as *pancytopenia* and, on bone marrow biopsy, fatty marrow replaces the red marrow. Pockets of HSC activity may still be present although the maturing cells often exhibit dysplastic morphology.

Methemoglobinemia occurs when the iron atom looses a single electron, referred to as *methemoglobin*. The iron in methemoglobin is in a higher spin state and can no longer form a reversible complex with molecular oxygen. Fortunately, normal RBCs can reduce methemoglobin by way of reductases that are present in the RBC.

TABLE 5.3 Increase Destruction of RBC: Hemolytic Anemia

Primary disorders of red cell membranes
Secondary disorders of red cell membranes
Immunohemolytic anemia
Hemolytic anemia caused by RBC infection
Heinz bodies hemolytic anemia–oxidative damage
Hemolytic anemia caused by glycolytic defects
Hemoglobinopathies
Anemia of splenomegaly
Hemolytic anemia incidental to generalized diseases

Source: Adapted from Jandl (1987), *Blood Textbook of Hematology.*

Neutropenia and Agranulocytosis

The decrease in the number of circulation neutrophils is referred to as *neutropenia* and occurs idiopathically, from a hereditary or congenital defect, following severe infection, immune-mediated, and from the use of specific drugs (e.g., antimicrobials, phenothiazines, anti-inflammatory). However, these are not always independent initiation events for the onset of neutropenia. For example, some neutropenic patients produce antineutrophil antibodies directed toward the destruction of neutrophils. Most cases of immune-mediated neutropenia are drug-induced events. On the other hand, neutropenia is labeled as idiopathic when none of the these hereditary or drug-induced events can be defined.

Neutropenia marked by a neutrophil count less than 500 cells/μl constitutes *agranulocytosis*. Agranulocytosis is a relative increase in peripheral blood cells lacking granules in their cytoplasm (e.g., lymphocytes). *Granulocytopenia* is a term used for the reduction of all cells in the granulocytic series. Since the differential WBC count indicates the

TABLE 5.4 Impaired Production of RBCs

Category	Functional Defect	Bone Marrow Morphology	Red Cell Morphology	Common Causes
Aplastic anemia	Disturbance in stem cell kinetics	Hypoplastic	Normocytic or macrocytic, normochromic	Chemicals, radiation, renal deficiency, carcinoma, idiopathic
Megaloblastic anemia	Impaired DNA synthesis	Hyperplastic, megaloblastic	Macrocytic, normochromic	Vitamin B_{12} deficiency, folate deficiency
Hypochromic anemia	Impaired hemoglobin synthesis	Hyperplastic, deficient hemoglobinization	Microcytic, hypochromic	Iron deficiency, anemia of chronic disease, thalassemia, sideroblastic disorders

Source: Adapted from Jandl (1987), *Blood Textbook of Hematology.*

percentage of leukocytes that are present in the peripheral blood, a decrease in granulated cells manifests as an increase in cells without granules. Thus, neutropenia and agranulocytosis often appear hand in hand. See also a review by Mavroudi and Papadaki (2012).

Thrombocytopenia

Thrombocytopenia refers to a decrease in the number of platelets circulating in the peripheral blood. The clinical signs of thrombocytopenia are skin discoloration such as reddish-purple blotches (i.e., purpura), large bruise-like patches (i.e., ecchymoses), and bleeding in small areas of the skin and mucous membranes (i.e., petechia), or spontaneous bleeding can occur as a result of a decreased number of platelets. The bone marrow analysis in thrombocytopenia shows either an increase or normal megakaryocyte counts, although immature megakaryocytes are present. These bone marrow findings represent a hallmark of increased demand for peripheral blood platelets. In contrast, a reduced production of thrombocytes is marked by small, pyknotic platelets in the peripheral blood and few megakaryocytes in the bone marrow. Severe thrombocytopenia can occur as an isolated incident, such as idiopathic thrombocytopenic purpura, or as part of a wider dysfunction, such as bone marrow failure. However, the term idiopathic is historical, in that the trigger for thrombocytopenia is most often uncovered in the modern-day clinical setting.

Drug-induced immune thrombocytopenia is a process initiated by an antibody response to a drug or a metabolite of the drug. The antibody binds to the thrombocyte and then the antibody–thrombocyte complex is recognized and cleared by macrophages. There are many pharmaceutical drugs that can initiate this phenomenon, but the most common triggers are analgesics/anti-inflammatory drugs and antibiotics.

Aplastic Anemia

Aplastic anemia deserves further discussion as this condition has been associated with numerous toxic agents and, in some cases, it has been considered as part of a linear disease progression toward acute myeloid leukemia (AML). Aplastic anemia is a rare bone marrow disease that has been recognized as a clinical entity for more than a century. The essential features of bone marrow failure associated with aplastic anemia are a peripheral blood pancytopenia accompanied by a hypocellular bone marrow in the absence of any congenital, neoplastic, or malabsorption disorder. The clinical presentation of aplastic anemia varies from an acute, life-threatening illness to a progressive, debilitating disease. Historically, the study of aplastic anemia was difficult due to a combination of several factors, including imprecision in the differential diagnosis and uncertainties regarding the pathogenesis of the disease. The fact that aplastic anemia is not the most frequent cause of pancytopenia noted in peripheral blood assays further compounds these issues. The incidence of aplastic anemia in the general population is extremely low (i.e., ~1/100,000 worldwide). Even for those individuals exposed to known etiological agents, the development of the disease is still relatively rare. Various etiological agents include numerous pharmaceuticals, viruses, ionizing radiation, and chemicals—most notably, benzene. All have all been implicated in the development of this disease.

TABLE 5.5 Subtypes of Myeloproliferative Neoplasms

Chronic myelogenous leukemia *BCR-ABL* positive (CML)
Chronic neutrophilic leukemia (CNL)
Polycythemia vera (PV)
Essential thrombocytopenia (ET)
Primary myelofibrosis
Chronic eosinophlic leukemia (CEL)
Mastocytosis
Myeloproliferative neoplasm-unclassifiable (MPN-U)

Source: Adapted from Swerdlow (2008).

Myeloproliferative Neoplasms

Myeloproliferative neoplasms are clonal HSC disorders characterized by proliferation of one or more myeloid lineage (i.e., granulocytes, erythrocytes, megakaryocytes, and mast cells). These neoplasms primarily arise in the fifth to seventh decade of life, although some of the disease subtypes have been reported in children (e.g., chronic myelogenous leukemia (CML) and essential thrombocytopenia). The incidence of all subtypes of myeloproliferative neoplasms combined is 6–10/100,000 individuals worldwide.

In the clinical setting, myeloproliferative neoplasms are characterized by hypercellularity of the bone marrow with effective hematopoietic maturation and, depending on the subtype, increase in the number of granulocytes, RBCs, and/or platelets in the peripheral blood (Table 5.5). Discussion of individual myeloproliferative subtypes is limited to the more commonly diagnosed diseases and/or diseases known to be associated with hematotoxins.

Chronic Myelogenous Leukemia Under the French–American–British (FAB) classification scheme, CML was considered in the same category as other myeloid leukemias. However, the classification of CML as a myeloproliferative neoplasm may be more appropriate considering the indolent nature of this disease, along with the increase in myeloid cells.

CML is a clonal disease thought to originate in a pluripotent bone marrow stem cell. Interestingly, the diagnostic criteria for CML is the presence of the Philadelphia chromosome, a balanced translocation of chromosome 9 with chromosome 22 [t(9;22)]. This translocation fuses the

ABL1 gene on chromosome 9 with the *BCR* gene on chromosome 22. The *BRC/ABL-1* fusion gene results in an abnormal fusion protein with increased tyrosine kinase activity, resulting in the constitutive activation of several cell signaling pathways. The defect in CML manifests as deregulation of proliferation, reduced adherence to the bone marrow stroma, and abnormal response to apoptotic stimuli. CML has an annual incidence of 1–2/100,000 individuals worldwide, and the median age at diagnosis is in the fifth or sixth decade of life. Aside from radiation exposure, factors predisposing individuals to CML are unknown and there does not appear to be a hereditary disposition.

Typically, CML is a bi- or triphasic disease with overlapping boundaries. The chronic phase (CP) is marked by an increase in minimally invasive myeloid cells that are mostly confined to the hematopoietic tissue, including the peripheral blood, bone marrow, spleen (i.e., splenomegaly), and occasionally the liver (i.e., hepatomegaly). The accelerated phase (AP) is largely referred to as the transformation phase where there is a persistent increase in the number of myeloid cells that are unresponsive to therapy. In the blast phase (BP) of the disease, myeloblasts compose at least 20% of the peripheral blood WBC count. In approximately 70% of the cases, the blast cells consist of myeloid cells, but it is also common to observe a combination of neutrophilic, eosinophilic, basophilic, monocytic, megakaryocytic, and erythroid blast cells. It is interesting to note that in 20–30% of the cases, the blast cell population also includes lymphoblasts. In addition, blast cell proliferation can infiltrate anywhere in the body, although it most commonly involves the skin, lymph nodes, bone, and central nervous system. Most patients are diagnosed in the CP, and they are often asymptomatic even though they exhibit an increased WBC count (i.e., median 100×10^9/l) due to an increase in the percentage of myelocytes and segmented neutrophils in the peripheral blood. The bone marrow is hypercellular, also due to an increase in granulocytic maturation.

The increased understanding of the characteristics associated with the *BCR/ABL* fusion protein and the resulting tyrosine kinase activity, as well as the bi- or triphasic nature of CML disease progression, has led to targeted drug development at the kinase binding domain. Imatinib is a first-generation drug that competes with ATP for kinase binding on the fusion protein. Imatinib binding results in the inhibition of phosphorylation of downstream substrates, thus disrupting the oncogenic signaling and effectively controlling the disease, especially if imatinib is used in the CP.

Polycythemia Vera PV is also classified as a myeloproliferative neoplasm characterized by an increase in the production of erythrocytes independent of normal mechanisms of erythropoiesis. Nearly all patients with PV exhibit a mutation in a kinase gene resulting in abnormal kinase activity. The clinical manifestation of the kinase mutation is proliferation in the erythroid, granulocytic, and megakaryocytic lineages, which is referred to as *panmyelosis*. As with other myeloproliferative neoplasms, there are three distinct phases of PV disease progression: a prepolycythemic phase, which is characterized by a mild erythrocytosis; an overt polycythemic phase, in which the red cell mass is significantly increased; and a spent or postpolycythemic phase associated with myelofibrosis, peripheral blood cytopenias, bone marrow failure, and extramedullary hematopoiesis. An additional feature associated with PV is a low incidence of disease progression to myelodysplastic syndromes (MDS) and/or AML.

The incidence of PV increases with age and varies between 0.7 and 2.6/100,000 individuals in European and North American populations. The etiology of PV is largely unknown although ionizing radiation and occupational exposures to toxins have been suggested as potential causes in some patients. Additionally, a genetic predisposition has been indicated in some families.

Primary Myelofibrosis Myelofibrosis is another subtype of myeloproliferative neoplasm. Interestingly, the proliferating lineages in primary myelofibrosis are predominantly megakaryocytes and granulocytes. Myelofibrosis is a stepwise disease beginning with bone marrow hypercellularity. As the disease progresses and becomes fully developed, myelofibrosis is characterized by the deposition of fibrous connective tissue such as reticulin and collagen within the bone marrow space as well as evidence of extramedullary hematopoiesis. The fibrotic stage of this disease is also characterized by leuko-erythroblastosis in the peripheral blood, which is marked by teardrop-shaped RBCs.

The incidence of primary myelofibrosis is approximately 0.5–1.5/100,000 persons annually and the incidence increases with age, most commonly occurring in the sixth or seventh decade of life. Myelofibrosis also occurs equally in men and women. Although it is rarely seen in children, cases of familial myelofibrosis have been reported in children and in some instances it appears to represent an autosomal recessive inherited condition. In other families, familial myelofibrosis presents at a later age of onset. In addition to a familial disposition and ionizing radiation, a single case of benzene exposure has been documented as a potential cause of myelofibrosis.

Myelodysplastic/Myeloproliferative Neoplasms

Myelodysplastic/myeloproliferative neoplasms (MDS/MPN) are clonal neoplasms that share characteristics of both MDS and myeloproliferative disorders (Table 5.6). Patients usually present with hypercellular bone marrow due to proliferation in one or more myeloid lineage. The combination of disease characteristics is also observed in some cases where proliferation of a cell lineage is successful but the cell morphology is dysplastic. In turn, peripheral blood

TABLE 5.6 Subtypes of Myelodysplastic/Myeloproliferative Neoplasms

Chronic myelomonocytic leukemia (CMML)
Atypical chronic myeloid leukemia BCR/ABL1 negative (aCML)
Juvenile myelomonocytic leukemia (JMML)
Myelodysplastic/myeloproliferative neoplasm—unclassifiable (MDS/MPN-U)

Source: Adapted from Swerdlow (2008).

cytopenias may result from ineffective lineage proliferation. It is important to note that as part of the criteria for diagnosis of MDS/MPN, the number of blast cells in both the blood and bone marrow is always less than 20%. Further discussion of individual myelodysplastic/myeloproliferative subtypes in this book chapter is limited to CMML as it is a hematopoietic disease that has been associated with toxins.

Chronic Myelomonocytic Leukemia CMML is a clonal hematopoietic malignancy that is heterogeneous in clinical appearance with characteristics of both MDS and MPN. CMML has been described as erythromonocytic leukemia, subacute myelomonocytic leukemia (AMML), subacute myeloid leukemia, and a smoldering leukemia. There has been debate regarding CMML's classification because of its variability in clinical presentation. Previous studies grouped CMML with the myeloproliferative neoplasm, CML, while others regarded CMML as a subtype of MDS. Although the 1976 FAB classification system formally recognized CMML as a distinct entity, FAB later subclassified CMML as an MDS-like disease (WBC $< 13 \times 10^9$/l) or a MPN-like disease (WBC $> 13 \times 10^9$/l). Gene expression analyses are consistent with the proliferative component in CMML, distinguishing it from MDS. Therefore, CMML was classified in the MDS/MPN category by WHO 2001 criteria and continues to be classified as MDS/MPN by the current 2008 WHO criteria.

The essential clinical and hematological features of CMML include an increase in circulating monocytes, with frequent descriptions of anemia and thrombocytopenia. In addition to peripheral blood abnormalities, patients with CMML frequently present with multilineage dysplasia and bone marrow hypercellularity. No characteristic cytogenetic abnormality has been described for CMML, although chromosome abnormalities have been reported in approximately 30% of CMML cases. These abnormalities include trisomy 8, monosomy 7, del(7q), and abnormalities involving chromosome 12p. The onset of CMML is usually subtle and occurs mostly in patients older than 50 years.

The differing clinical presentations of CMML and the historical classification debate have complicated epidemiological investigations of CMML. As a result, reliable incidence data are sparse. The etiology of CMML is also unknown, although ionizing radiation as well as occupational and environmental exposures to carcinogenic agents have been suggested as possible causes of this disease.

Myelodysplastic Syndromes

MDS present as a heterogeneous set of neoplastic diseases that are characterized by cytopenias in the peripheral blood, dysplasia in the bone marrow, and progressive failure of hematopoiesis in one or more of the myeloid cell lineages. MDS are thought to originate in HSC and a diagnosis of MDS offers an increased risk for the development of certain types of AML. Although progression of MDS to AML can occur as a natural course of the disease, the diagnostic distinction between MDS and AML is determined by the threshold percentage of myeloblastic cells observed in the peripheral blood and bone marrow: greater than 20% myeloblasts meets the criteria for a diagnosis of AML and less than 20% myeloblasts meets the criteria for a diagnosis of MDS.

Incidence of *de novo* MDS increases with age with a male predominance. The median age at onset is in the seventh decade of life and the incidence rises in the seventh decade from 3–5/100,000 to 20/100,000 persons. In addition to age, other possible causes of MDS include chemotherapy, radiation exposure, exposure to benzene, agricultural chemicals and solvents, cigarette smoking, and a family history of hematopoietic neoplasms or inherited hematopoietic disorders.

Historically, epidemiological reports for hematopoietic neoplastic disease have not defined MDS or MDS subtypes, which has made it extremely difficult to attribute specific subtypes to causes of disease in the epidemiology literature. The changing definition of the various subtypes during the past 30 years has further complicated matters. For example, MDS was often referred to as preleukemia, a preleukemic state, or subacute leukemia, none of which distinguished MDS from other hematopoietic diseases. As such, the historic reports often included either a description of MDS along with a diagnosis of AML or MDS categorized with other chronic myeloproliferative diseases (e.g., myeloid fibrosis).

In 1976, the FAB cooperative proposed a classification scheme for "dysmyelopoietic syndromes," which referred to two subtypes of bone marrow conditions wherein the bone marrow exhibited hypercellularity with minimal blast cell numbers. The subtypes included refractory anemia with excess blasts (RAEB) and CMML. In 1982, the FAB system began referring to the syndromes as "myelodysplastic" and proposed a new classification scheme. Following FAB's review of the morphological features of 80 separate cases 5 distinct subtypes were defined. These subtypes included refractory anemia (RA), RA with ring sideroblasts sideroblasts (RARS), RAEB, RAEB in transformation (RAEB-t), and CMML. In 1997, the WHO appointed a new committee that revised and updated the FAB classification scheme. At that time, the committee took into account blast cell

TABLE 5.7 Subtypes of Myelodysplastic Syndromes

Refractory cytopenia with unilineage dysplasia (RCUD)
Refractory anemia (RA)
Refractory neutropenia (RN)
Refractory thrombocytopenia (RT)
Refractory cytopenia with multilineage dysplasia (RCMD)
Refractory anemia with ring sideroblasts (RARS)
Refractory anemia with excess blasts-1 (RAEB-1)
Refractory anemia with excess blasts-2 (RAEB-2)
Myelodysplastic syndrome—unclassifiable (MDS-U)
MDS associated with isolated del(5q)
Childhood myelodysplastic syndrome

Source: Adapted from Swerdlow (2008).

percentage, dysplastic features, prognostic features, and therapeutic outcomes in defining the criteria for diagnosis.

Today, advances in tumor biology and diagnostic tools used to identify histopathology and genetic features allow WHO to set forth uniform diagnostic criteria for MDS subtypes (Table 5.7). The MDS subtypes are morphologically categorized by their varying degree of blast cells, lineage dysplasia, and the presence of erythroid precursor cells with an accumulation of iron in the mitochondria (i.e., ring sideroblasts).

2001 WHO Criteria for MDS Subtypes The initial WHO criterion was published in 2001 and several distinct changes were noted from the 1982 FAB classification. First, RAEB-t was eliminated as an MDS subtype and bone marrow with greater than 20% blast cells was considered as an AML. Second, the categories of RAEB-1 and RAEB-2 were created with blast cell percentage used as a subclassification distinction 5–9% and 10–19%, respectively. Third, CMML was eliminated as an MDS subtype and a new category, MDS/MPN, was created, which includes CMML. Finally, three additional MDS subtypes were created that encompassed cytopenias with multilineage dysplasias (RCMD), MDS with no classifiable features (MDS-u), and the 5q minus syndrome (i.e., distinct from 5q − chromosomal abnormality) (Jaffe, 2001).

2008 WHO Criteria for MDS Subtypes The latest WHO revision of the MDS classification scheme was published in 2008 and reaffirmed that the majority of MDS subtypes were distinguished by the observed percentage of bone marrow blast cells, the presence or absence of ring sideroblasts, and the evidence of dysplasia. The 2008 WHO criteria are currently used for the diagnosis of hematopoietic diseases. The classification of MDS continues to evolve, however, and will likely be refined as we learn more about the biology and etiology of these diseases (Swerdlow, 2008).

Refractory Cytopenia with Unilineage Dysplasia In the WHO 2008 classification scheme, refractory cytopenia with unilineage dysplasia (RCUD) represents 10–20% of all cases and the median age of onset is 65–70 years. RCUD includes RA, refractory neutropenia (RN), and refractory thrompocytopenia (RT), defined as a single cytopenia in the peripheral blood and a single lineage dysplasia in the bone marrow. RCUD is an MDS subtype that is most commonly associated with erythroid dysplasia, with RA representing the majority of RCUD cases. Bone marrow findings for RCUD include less than 5% myeloblasts and less than 15% of the erythroid precursors are ring sideroblasts. RCUD has a low incidence of transformation to AML.

Refractory Anemia with Ring Sideroblasts RARS also represents MDS, a subtype most commonly associated with erythroid dysplasia. RARS are defined by anemia in the peripheral blood and erythroid dysplasia in the bone marrow. Additional bone marrow findings for the classification of RARS include the requirement for greater than 15% of the erythroid precursor cells to be ring sideroblasts and the observation of less than 5% myeloblasts. Evidence suggests that RARS is a clonal stem cell defect that manifests as abnormal iron metabolism in erythroid cells and ineffective erythropoiesis. Like RCUD, RARS has a low incidence of transformation to AML.

Refractory Cytopenia with Multilineage Dysplasia Refractory cytopenia with multilineage dysplasia (RCMD) is represented by one or more cytopenias in the peripheral blood and dysplasia in at least two myeloid lineages in the bone marrow. The blast cells are represented at less than 5% in the bone marrow. Interestingly, cytogentic abnormalities such as trisomy 8, monosomy 7, del(7q), monosomy 5, del(5q), and del(20) are noted in up to 50% of patients with RCMD. Further, RCMD accounts for approximately 30% of all diagnosed MDS. There is an increased incidence of RCMD with increasing age, a median onset of 70 years, and a male predominance.

Refractory Anemia with Excess Blasts (RAEB-1 and RAEB-2) WHO 2008 defines two subcategories of refractory anemia with excess blasts (RAEB-1 and RAEB-2) based on the percentage of blast cells observed in the peripheral blood and the bone marrow. Cases of RAEB-1 are defined by 5–9% blasts in the bone marrow. In contrast, cases of RAEB-2 are defined by 10–19 blast cells in the bone marrow. When cases of RAEB show an increase in blast cells that are greater than 20% in the bone marrow, then a diagnosis of AML is considered. Both RAEB-1 and RAEB-2 are characterized by cytopenias in the peripheral blood and unilineage or multilineage dysplasia in the bone marrow. RAEB is seen more often in individuals older than 50 years and accounts for 40% of diagnosed MDS cases.

MDS with del(5q) It is noteworthy that MDS with an isolated 5q deletion is classified as an independent MDS

subtype under WHO 2001 and WHO 2008 criteria. This subtype is unique in that cells exhibit a single del(5q) cytogenetic abnormality and the patient may present with an increased number of thrombocytes (thrombocytosis) instead of thrombocytopenia in the peripheral blood. Additional features characteristic of MDS del(5q) include anemia with or without other cytopenias in the peripheral blood and blast cells observed at less than 5% in the bone marrow. MDS with an isolated 5q deletion occurs more often in women with a median onset age of 67 years. Some suggest that a defect in ribosomal protein function is causally associated with this disease.

Myelodysplastic Syndrome—Unclassifiable (MDS-U) Myelodysplastic syndrome—unclassifiable (MDS-U) represents a clear case of bone marrow dysplasia, but it does not fit under any other MDS subtype classification. There are two indications for classification in the MDS-U category. The first is an MDS case with lineage dysplasia in one or more of the myeloid lineages that is accompanied by pancytopenia, RCUD, or RCMD subtype that exhibit less than 1% blast cells in the peripheral blood. The second is an MDS case represented by bone marrow dysplasia in one or more lineage with fewer than 1% blast cells in the peripheral blood and less than 5% blasts in the bone marrow that is accompanied by a cytogenetic abnormality. The incidence of MDS-U is unknown.

As previously mentioned, the WHO classification schema for clonal hematopoietic neoplasms are a "work in progress" and, as our knowledge increases, MDS will likely be further defined. For additional information please see Bennett (2005); Jaffe (2001); and Swerdlow (2008).

Acute Myeloid Leukemia

According to the 2008 WHO classification of hematopoietic tumors, a disease must be described, defined, and named before it can be diagnosed, treated, and studied. The 2008 WHO classification of tumors describe multiple subtypes of AML. The WHO criteria incorporates known clinical features, and, as discussed in Hematotoxicity Section 5.2, morphology, cytochemistry, immunophenotype, genetics, epidemiology, etiology, and prognostic features. Discussion in this section is limited to the leukemic subtypes listed in Table 5.8.

TABLE 5.8 Subtypes of Acute Myeloid Leukemia

AML with recurrent genetic abnormalities
AML with t(8;21)
AML with inv(16) or t(16;16)
APL with t(15;17)
AML with t(9;11)
AML with gene mutations
AML with myelodysplastic-related changes
Therapy-related t-AML/t-MDS
Alkylating agents
Radiation therapy
Topoisomerase II inhibitors
AML not otherwise specified
AML with minimal differentiation
AML without maturation
AML with maturation
Acute myelomonocytic leukemia
Acute monoblastic monocytic leukemia
Acute erythroid leukemia
Acute megakaryocytic leukemia
Acute basophilic leukemia

Source: Adapted from Swerdlow (2008).

AML with Recurrent Genetic Abnormalities The balanced translocations described in this classification subtype are actually structural chromosome rearrangements that create functional fusion proteins and these functional fusion proteins are a requirement for leukemic transformation.

AML with t(8;21)(q22;q22) The balanced translocation created by t(8;21)(q22;q22) is the *RUNX1-RUNX1T1* gene, which encodes for the core binding factor alpha (CBFα) subunit, a transcription factor essential for hematopoiesis. It has been suggested that the fusion protein likely results in transcriptional repression of normal *RUNX1* target genes as the phenotype associated with AML with t(8;21)(q22;q22) shows maturation in the granulocytic lineage. AML with t(8;21)(q22;q22) occurs in approximately 5% of AML cases and is usually associated with a good response to chemotherapy.

AML with inv(16)(p13.1q22) or t(16;16)(p13.1:q22) The translocation created by the common inv(16)(p13.1q22) or much less common t(16;16)(p13.1:q22) results in the fusion of the core binding factor beta subunit (*CBFb*) gene located at 16q22 and the smooth muscle myosin heavy chain (*MYH11*) gene. The *CBFB-MYH11* fusion gene manifests as an AML with monocytic and granulocytic differentiation and an abnormal eosinophil component in the bone marrow. AML with inv(16)(p13.1q22) or t(16;16)(p13.1:q22) represents 5–8% of AML cases and occurs predominately in younger patients.

APL with t(15;17)(q22;q12) The balanced translocation in acute promyelocytic leukemia (APL) is created by the fusion of retinoic acid receptor alpha (*RARA*) gene located at the chromosome 17q12 loci and the nuclear regulatory factor (*NRF*) gene located at chromosome 15q22. NRF loci are also referred to as the *promeylocytic leukemia* gene (PML). The prognosis for patients with APL t(15;17)(q22;q12) treated with all-trans retinoic acid is favorable as RARA protein has an affinity for all trans retinoic acid, which acts as a differentiating agent. APL with t(15;17)(q22;q12) is found

in 5–8% of AML cases. APL can occur at any age although it is seen most prominently at middle age.

AML with t(9;11)(p22;q23) The translocation associated with t(9;11)(p22;q23) involves the *MLL* fusion gene located at 11q23. The *MLL* gene loci code for a histone methyltransferase protein that regulates the transcription process by chromatin remodeling. The t(9;11)(p22;q23) *MLLT3-MLL* is usually associated with monocytic features and is the most common MLL translocation in AML. It may occur at any age although it is more common in children. It accounts for 9–12% of childhood AML and 2% of adult AML.

AML with gene mutations AML with gene mutations occur in AML patients with a normal karyotype. The most common mutations include the fms-related kinase 3 (*FLT3*), nucleophosmin 1 (*NPM1*), and mutations in the gene for CCAAT/enhancer binding protein-alpha (*CEBPα*). NPM1 mutations occur in approximately one-third of all AMLs, whereas CEBPα mutations occur in 6–15% of all AML. Additional mutations that occur alone or in combination with *FLT3*, *NPM1*, and *CEBPα* genes include *KIT*, *MLL*, *WT1*, *NRAS*, and *KRAS*.

AML with Myelodysplastic-Related Changes AML with myelodysplastic-related changes is synonymous with acute myeloid leukemia with multilineage dysplasia (AML-MD). Patients with AML-MD usually present with dysplasia in two or more cell lineages in the bone marrow and pancytopenia in the peripheral blood. Blast cell counts greater than 20% are noted in the peripheral blood and the bone marrow. The criteria for diagnosis of AML-MD include a previous history of MDS and MDS-related cytogenetic abnormalities. Genetic abnormalities that occur most commonly with AML-MD include the loss or gain of chromosome segments such as the unbalanced chromosome abnormalities −7/del(7q) and −5/del(5q). Less common translocations include t(3;5)(q25;q34), inv(3)(q21q26.2), t(3;3)(q21q26.2), or t(6;9)(p23;q34). In order to meet the criteria for diagnosis of AML-MD there must be an absence of a prior history of therapy with cytotoxic agents for an unrelated disease.

AML-MD occurs mainly in older patients and rarely in children. AML-MD represents 24–35% of all AML cases. The prognosis for this disease is low as compared to other types of AML.

Therapy-Related (t) t-AML/t-MDS This category of AML includes tAML, tMDS, and tMDS/MPN occurring as a complication of cytotoxic and/or radiation therapy administered following the diagnosis of a primary neoplasm. The actual classification as tMDS or tAML is a function of the number of blast cells present. MPN alone is not included in this category as evolution of the disease can result in AML.

TABLE 5.9 Two Categories of Therapy-Related Hematopoietic Neoplasms and Corresponding Chromosome Abnormalities

Alkylating agents/ionizing radiation
Abnormalities of chromosome 5
Abnormalities of chromosome 7
Topoisomerase II inhibitors
11q23 including t(9;11)(p22;q23); t(11;19)(q23;p31
21q22 including t(8;21)(q22;q22) and t(3;21)(q26.2;q22)
t(15;17)(q22;q12)
inv(16)(p13q22)

Source: Adapted from Swerdlow (2008).

Therapy-related AML accounts for 10–20% of all AML cases. The incidence varies according to the underlying disease and treatment. Any age group can be at risk, although tAML following alkylating agents or radiation therapy appears to increase with age, whereas the risk for tAML following topoisomerase II inhibitors is similar for all ages.

Therapy-related AML is thought to occur through mutational events induced by the cytotoxic agents used during therapy. Some people may be genetically predisposed to therapy-related complications or have defects in DNA repair mechanisms. For most cases, however, the entire pathway toward AML is unclear.

There are two subsets of therapy-related hematopoietic neoplasms (Table 5.9). The most common subset of tAML/tMDS and tAML/tMDS/MPN occurs 5–10 years post-treatment with alkylating agents and/or ionizing radiation. Patients initially present with tMDS in the form of bone marrow failure along with peripheral blood cytopenias. However, a minority of cases treated with alkylating agents and/or ionizing radiation present with tMDS/MPN or overt tAML. These cases are most commonly associated with the unbalanced loss of genetic material involving chromosomes 5 and 7.

The second category of therapy-related hematopoietic neoplasms include tAML/tMDS and tAML/tMDS/MPN and occurs in 20–30% of the patients with a latency of 1–5 years following cytotoxic therapy with topoisomerase II inhibitors. Most patients do not present with tMDS but instead present with overt tAML. The topoisomerase II–related tAML is often associated with balanced translocations such as 11q23, 21q22, t(15;17), and inv(16). Please refer to Larson and LeBeau (2005) for further information on therapy-related leukemias.

AML Not Otherwise Specified This category includes those cases of AML that do not fit in any other more specific category and is largely based on the original FAB classification schema.

AML with Minimal Differentiation AML with minimal differentiation represents an AML with no evidence of a myeloid lineage differentiation by morphology or cytochemistry.

However, immunological markers such as CD34, CD38, and, in some 60% of the cases, CD33 are indicative of myeloid phenotype. Cases of AML with minimal differentiation represent less than 5% of AML cases and are most common in either infants or older adults. Patients present with bone marrow failure and cytopenias such as anemia, thrombocytopenia, and neutropenia. In addition, there is often leukocytosis and a large number of blast cells.

AML without Maturation AML without maturation is evidenced by a large number of blast cells, greater than 90%, within the bone marrow with no evidence of maturation to neutrophils. This disease comprises 5–10% of AML cases and the median age of onset is 46 years. Again, patients present with bone marrow failure, anemia, thrombocytopenia, and neutropenias, as well as leukocytosis with a large number of blast cells.

AML with Maturation AML with maturation is characterized by greater than 10% of maturing neutrophils present in the peripheral blood although the bone marrow consists of less than 20% cells of the monocytic lineage. This disease represents 10% of all AML cases and it can occur at any age group. Patients often present with anemia, thrombocytopenia, and neutropenia. Leukocyte and blast cell counts are variable.

Acute Myelomonocytic Leukemia The characteristics associated with AMML include the proliferation of both neutrophilic and monocytic precursor cells, which make up 20% of the bone marrow. An increased number of monocytes may also be observed in the peripheral blood. AMML comprises 5–10% of all AML cases and the median age at onset is 50 years. There is a slight male predominance. Patients usually present with fever, fatigue, anemia, and thrombocytopenia.

Acute Monoblastic and Monocytic Leukemia In acute monoblastic and monocytic leukemia, 80% or more of the leukemic cells are of monocytic lineage with neutrophils as a minor component. This disease represents less than 5% of the AML cases and is most common in young patients. Bleeding disorders are common and extramedullary lesions and masses may occur with infiltration into the cutaneous and/or gingival tissue, as well as infiltration into the central nervous system.

Acute Erythroid Leukemia Acute erythroid leukemia is characterized by leukemia with a predominant erythroid lineage. There are two subtypes of erythroid leukemia distinguished by the presence or absence of a granulocytic component. The first subtype is erythroleukemia, characterized by 50% of the bone marrow composed of erythroid precursors. The second subtype is defined as pure erythroid leukemia, which represents the neoplastic proliferation of undifferentiated committed erythroid precursor cells in greater than 80% of the bone marrow cells with no evidence of a myeloblastic component. Erythroleukemia represents less than 5% of AML cases and predominately occurs in adults. Pure erythroid leukemia is rare although it can occur at any age, including childhood.

Acute Megakaryoblastic Leukemia Acute megakaryoblastic leukemia occurs when 50% of the blast cell population is composed of the megakaryocytic lineage. This disease occurs in less than 5% of all AML cases, although it is observed in both adults and children. Patients with acute megakaryoblastic leukemia present with thrombocytopenia, though some cases may present with thrombocytosis. Dysplastic features may be present in neutrophils, erythroid precursors, platelets, and megakaryocytes. Splenomegaly and hepatomegaly are also common.

Acute Basophilic Leukemia In acute basophilic leukemia the primary differentiated lineage is basophils. This disease is quite rare, present in less than 1% of all AML cases. Patients with acute basophilic leukemia may present with bone marrow failure, cutaneous involvement, organomegaly, lytic lesions, and symptoms related to hyperhistamine release.

5.4 HEMATOTOXINS: KNOWN HEMATOTOXIC AGENTS

This section correlates agents with known hematotoxic outcomes and a series of tables are used to present this information (Tables 5.10–5.20). For the majority of hematotoxins listed in the tables, the mechanisms associated with

TABLE 5.10 Agents Associated with Sideroblastic Anemia

Ethanol	Chloramphenicol
Isoniazid	Cooper chelation
Pyrazinamide	Zinc intoxication
Cycloserine	Lead intoxication

Source: Adapted from Casarett & Doull's 7th Edition, Chapter 11.

TABLE 5.11 Agents that Cause Megaloblastic Anemia

B_{12} Deficiency	Folate Deficiency
Paraminosalicylic acid	Phenytoin
Colchicine	Primidone
Neomycin	Carbamazepine
Ethanol	Phenobarbitol
Omeprazole	Sulfasalazine
Zidovudine	Cholestyramine
	Triamerine
	Antimetabolites

Source: Adapted from Casarett & Doull's 7th Edition, Chapter 11.

TABLE 5.12 Agents Known to Cause Aplastic Anemia

Chloramphenicol	Organic arsenicals	Quinacrine
Methylphenylethylhydantoin	Trimethadione	Phenylbutazone
Gold	Streptomycin	Benzene
Penicillin	Allopurinol	Tetracycline
Methicillin	Sulfonamides	Chlortetracycline
Sulfisoxazole	Sulfamethoxpyridazine	Amphotericin B
Mefloquine	Ethouximide	Felbamate
Carbimazole	Methylmercaptoimidazole	Potassium perchlorate
Propylthiouracil	Tolbutamide	Pyrimethamine
Chlorpropamide	Carbutamide	Tripelennamine
Indomethacin	Carbamazepine	Diclofenac
Meprobamate	Chloropromazine	Chlordiaepoxide
Mepazine	Chlorphenothane	Parathion
Thiocyanate	Methazolamide	Dinitrophenol
Bismuth	Mercury	Chlordane
Carbon tetrachloride	Cimetidine	Metolazone
Azidothymidine	Ticlopidine	Isoniazid
Trifluoperazine	D-penicillamine	

Source: Adapted from Casarett & Doull's 7th Edition, Chapter 11.

TABLE 5.13 Agents Associated with Methemoglobinemia

Therapeutic Agents	Environmental Agents
Benzocaine	Nitrites
Lidocaine	Nitrates
Prilocaine	Nitrobenzenes
Dapsone	Aniline dyes and derivatives
Amyl nitrate	Butyl nitrite
Isobutyl nitrite	Potassium chlorate
Nitroglycerine	Gasoline additives
Primaquine	Aminobenzenes
Sulfonamide	Nitrotoluenes
Phenacetin	Trinitrotoluene
Nitric oxide	Nitroethane
Phenazopyridine	Ortho-toluidine
Metoclopramide	Betanaphthol disulfonate
Flutamide	
Silver nitrate	
Quinones	
Methylene Blue	

Source: Adapted from Casarett and Doull's 7th Edition, Chapter 11.

TABLE 5.14 Agents that Cause Oxidative Damage to Erythrocytes

Acetanilide	Phenylhydrazine
Naphthalene	Nitrobenzene
Nitrofurantoin	Phenacetin
Sulfamethoxypyridazine	Phenol
Aminosalicytic acid	Hydroxylamine
Sodium sulfoxone	Methylene blue
Dapsone	Toluidine blue
Phenazopyridine	Furazolidone
Primaquine	Nalidixic acid
Chlorates	Sulfanilamide
Sulfasalazine	

Source: Adapted from Casarett and Doull's 7th Edition, Chapter 11.

TABLE 5.15 Agents that Induce Agranulocytosis

Chlorpromazine
Carbamazepine
Clozapine
Phenothiazines
Aminopyrine

Source: Adapted from Casarett and Doull's 7th Edition, Chapter 11.

TABLE 5.16 Agents that Cause Immune and Nonimmune Neutropenia

Drug associated with antibodies	Drugs not associated with antibodies
Aminopyrine	Rifampicin
Propylthiouracil	Ethambutol
Ampicillin	Allopurinol
Metiamide	Phenothiazines/CPZ
Dicloxacillin	Flurazepam
Phenytoin	Hydrochlorothiazide
Aprindine	
Azulfidine	
Chlorpropamide	
Phenothiazines/CPZ	
Procainamide	
Nafcillin	
Tolbutamide	
Lidocaine	
Methimazole	
Levamisole	
Gold	
Quinidine	
Clozapine	

Source: Adapted from Casarett and Doull's 7th Edition, Chapter 11.

TABLE 5.17 Agents that Induce Immune-Mediated Thrombocytopenia

Penicillin
Quinidine
Abciximab
Gold
Heparin
Ticlopidine
Mitomycin

Source: Adapted from Casarett and Doull's 7th Edition, Chapter 11.

initiation and subsequent progression of hematotoxicity is largely unknown. However, for some of these agents, classification into multiple mechanistic categories has been possible. For detailed information on mechanism of drug interaction please refer to Goodman and Gilman, 12th edition. For further information on mechanistic categories of hematopoietic disease please refer to Casarett and Doull's, 7th edition. For information on International Association of Research on Cancer (IARC) classification of agents that induce AML please see IARC Monograph 100 F (2012).

TABLE 5.18 Agents that Inhibit Coagulation Factors

Thrombin	Factor V	Factor VIII	Factor XIII	Von Willebrand Factor
Topical bovine thrombin	Topical bovine thrombin	Penicillin	Isoniazid	Ciprofloxacin
Fibrin glue	Penicillin	Ampicillin	Procainamide	Hydroxyethyl starch
	Gentamicin	Chloramphenicol	Penicillin	Valproic acid
	Cephalosporin	Phenytoin	Phenytoin	Griseofulvin
	Streptomycin	Methyldopa	Practolol	Tetracycline
		Nitrofurazone		Pesticides
		Phenylbutazone		

Source: Adapted from Casarett and Doull's 7th Edition, Chapter 11.

TABLE 5.19 Cytotoxic Agents Implicated in Therapy-Related Hematopoietic Neoplasms

Alkylating agents	Topoisomerase II inhibitors	Antimetabolites	Antitubulin agents
Melphalan	Etoposide	Thiopurines	Vincristine
Cyclophosphamide	Teniposide	Mycophenolate	Vinblastine
Nitrogen mustard	Doxorubicin	Fludarabine	Vindesine
Chlorabucil	Mitoxantrone		Paclitaxel
Busulfan	Amsacrine		Docetaxel
Carboplatin	Actinomycin		
Cisplatin			
Dicarbazine			
Procarbazine			
Carmustine			
Mitomycin C			
Thiotepa			

Source: Adapted from Swerdlow (2008).

TABLE 5.20 Agents Identified By IARC as Potential Leukemogens

Agent	Target Organ	Evidence of Genotoxicity
1,3,-Butadiene	Hemato-lymphatic organs	Strong
Benzene	Acute myeloid leukemia	Strong
Formaldehyde	Leukemia	Moderate

Source: Adapted from IARC Special Report (2009).

REFERENCES AND SUGGESTED READING

Bennett JM, Komrokji RS. The myelodysplastic syndromes: diagnosis, molecular biology and risk assessment. *Hematology* 2005;10 (Suppl 1):258–269.

Brunton LL. *Goodman & Gilman's The Pharmacological Basis of Therapeutics*. 12th ed. New York, Chicago, San Francisco: McGraw Hill Medical; 2011.

Greer JP, Foerster J, Rodgers GM, Paraskevas F, Glader B, Arber DA, Means RT. *Wintrobe's Clinical Zematology*. Volume 2. 12th ed. Philadelphia: Lippincott, Williams & Wilkins; 2008.

IARC. Monograph 100F Benzene, Formaldehyde, and 1,3 butadiene; 2012.

Jaffe ES, Harris NL, Stein H, Vardiman JW. *World Health Organization Classification of Tumours Pathology & Genetics Tumours of Haematopoietic and Lymphoid Tissues*. Lyon: IARC Press; 2001.

Jandl, JH. *Blood: A Textbook of Hematology*. 2nd ed. Chapter 7. Philadelphia: Lippincott, Williams & Wilkins; 1996.

Klaassen CD. *Casarett & Doull's Toxicology, The Basic Science of Poisons*. 7th ed. Chapter 11: Toxic Response of the Blood. New York: McGraw Hill; 2008.

Kourelis TV, Boruchov A, Hull D, Polio J, Scholes J, Moustakakis M, Bilgrami SF. Acute myeloid leukemia following solid organ transplantation: case report and comprehensive review. *Conn Med* 2012;76(3):151–154.

Larson RA, Le Beau MM. Therapy-related myeloid leukemia: a model for leukemogenesis in humans. *Chem Biol Interact* 2005;153–154:187–195.

Longo DL. *Harrison's Hematology and Oncology (Harrison's Specialties)*. New York, Chicago, San Francisco: McGraw Hill Companies; 2010.

Longo DL, Kasper DL, Jameson JL, Fauci AS, Hauser SL, Loscalzo J. *Harrison's Principles of Internal Medicine*. Volumes 1 and 2. 18th ed. New York, Chicago, San Francisco: McGraw Hill; 2008.

Mavroudi I, Papadaki HA. Genetic associations in acquired immune-mediated bone marrow failure syndromes: insights in aplastic anemia and chronic idiopathic neutropenia. *Clin Dev Immuno*l 2012;2012:123789.

Myers T. *Mosby's Medical Dictionary*. 8th ed. St. Louis: Elsevier Publishing; 2009.

Nissen-Druey C, Tichelli A, Meyer-Monard S. *Human Hematopoietic Colonies in Health and Disease*. Basel: S Karger; 2005. Re-Print from *Acta Haematologica* 113(1).

Pagana K, Pagana T. *Mosby's Diagnostic and Laboratory Test Reference*. 6th ed. St. Louis: Elsevier Publishing; 2010.

Renstrom J, Kroger M, Peschel C, Oostendorp RAJ. How the niche regulates hematopoietic stem cells. Chemico-Biol Interact 2010;184:7–15.

Ruffalo RR, Poste G, Metcalf RR. *Cell Cycle Regulation: New Horizons in Therapeutics*. Amsterdam: Overseas Publishers Association; 1997.

Shaffer LG, Slovak ML, Campbell LJ, editors. *ISCN: An International System for Human Cytogenetic Nomenclature*. Basel: S Karger; 2009.

Shapiro HM. *Practical Flow Cytometry*. 4th ed. Hoboken: John Wiley & Son, Inc.; 2003.

Shapiro HM. The cytometric future: it ain't necessarily flow. *Methods Mol Biol* 2011;699:471–482.

Sornberger J. *Dreams and Due Diligence: Till and McCulloch's Stem Cell Discover and Legacy*. Toronto, Buffalo, London: University of Toronto Press; 2011.

Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri SA, Stein H, Thiele J, Vardiman JW. *World Health Organization (WHO) Classification of Tumours of Haematopoietic and Lymphoid Diseases*. Lyon: IARC Press; 2008.

Wan TSK, Ma ESK. Molecular genetics: an indispensable tool for cancer diagnosis. Chang Gung Med J 2012;35(2):96–109.

6

HEPATOTOXICITY: TOXIC EFFECTS ON THE LIVER

ROBERT C. JAMES AND STEPHEN M. ROBERTS

The liver is one of the most important target organs for drug and chemical toxicity. The potential for hepatoxicity is often the principal concern in setting occupational and environmental standards for chemical exposure, and the discovery of hepatotoxicity is the most common cause for a drug to be removed from the market.

This chapter will familiarize the reader with

- The basis of liver injury
- Normal liver functions
- The role the liver plays in certain chemical-induced toxicities
- Types of liver injury
- Evaluation of liver injury
- Specific chemicals that are hepatotoxic

6.1 PHYSIOLOGICAL AND MORPHOLOGICAL BASES OF LIVER INJURY

Physiological Considerations

The liver is the largest organ in the body, accounting for about 5% of the total body mass. It is often the target organ of chemical-induced tissue injury, a fact recognized for over 100 years. While the chemicals toxic to the liver and the mechanisms of their toxicity are numerous and varied, several basic factors underlie the liver's susceptibility to chemical attack.

First, the liver maintains a unique position within the circulatory system. As Figure 6.1 shows, the liver effectively "filters" the blood coming from the gastrointestinal tract and abdominal space before this blood is pumped through the lungs and into the general circulation. This unique position in the circulatory system aids the liver in its normal functions, which include (i) carbohydrate storage and metabolism; (ii) metabolism of hormones, endogenous wastes, and foreign chemicals; (iii) synthesis of blood proteins; (iv) urea formation; (v) metabolism of fats; and (vi) bile formation. When drugs or chemicals are absorbed from the gastrointestinal tract, virtually all of the absorbed dose must pass through the liver before being distributed through the bloodstream to the rest of the body. Once a chemical reaches the general circulation, regardless of the route of absorption, it is still subject to extraction and metabolism by the liver. The liver receives nearly 30% of cardiac output and, at any given time, 10–15% of total blood volume is present in the liver. Consequently, it is difficult for any drug or chemical to escape contact with the liver, an important factor in the role of the liver in removing foreign chemicals.

The liver's prominence causes it to have increased vulnerability to toxic attack. The liver can particularly affect, or be affected by, chemicals ingested orally or administered intraperitoneally (i.e., into the abdominal cavity) because it is the first organ perfused by blood containing the chemical. As discussed in Chapter 2, rapid and extensive removal of the chemical by the liver can drastically reduce the amount of drug reaching the general circulation—termed the *first-pass effect.*

Being the first organ encountered by a drug or chemical after absorption from the gastrointestinal tract or peritoneal space also means that the liver often sees potential toxicants at their highest concentrations. The same drug or chemical at the same dose absorbed from the lungs or through the skin, for example, may be less toxic to the liver because the concentrations in blood reaching the liver are lower, from both dilution and distribution to other organs and tissues.

Principles of Toxicology: Environmental and Industrial Applications, Third Edition. Edited by Stephen M. Roberts, Robert C. James and Phillip L. Williams.

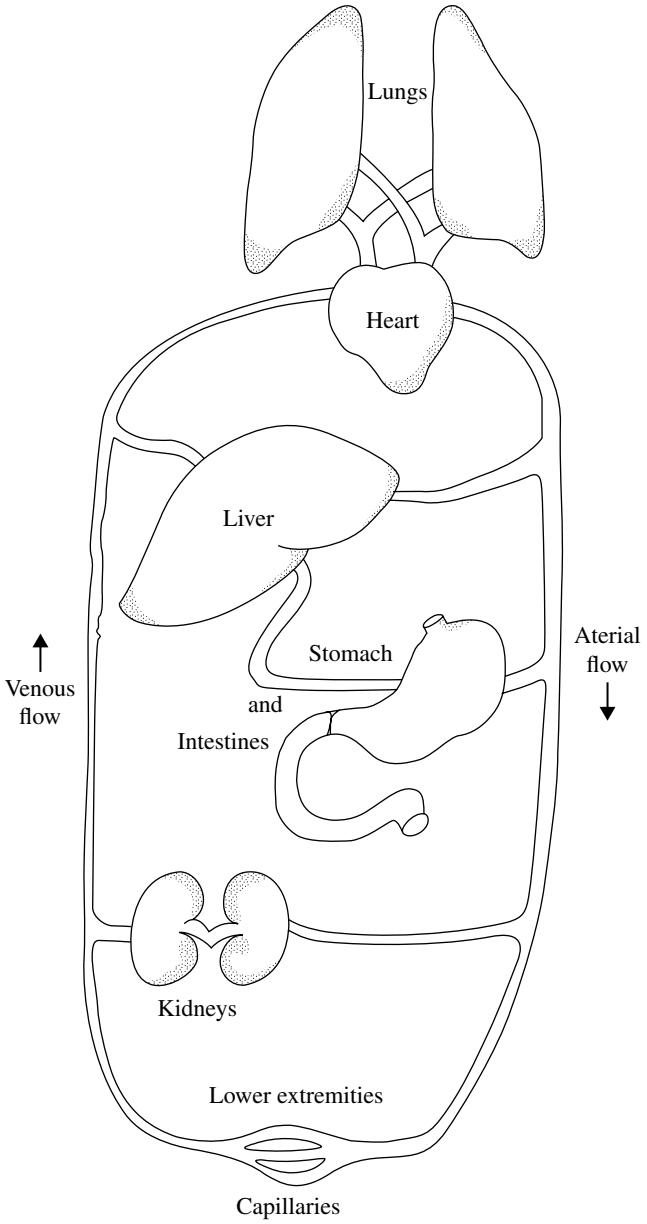


FIGURE 6.1 The liver maintains a unique position within the circulatory system.

A second reason for the susceptibility of the liver to chemical attack is that it is the primary organ for the biotransformation of chemicals within the body. As discussed in Chapter 3, the desired net outcome of the biotransformation process is generally to alter the chemical in such a way that it is (i) no longer biologically active within the body and (ii) more polar and water-soluble and, consequently, more easily excreted from the body. Thus, in most instances, the liver acts as a *detoxification* organ. It lowers the biological activity and blood concentrations of a chemical that might otherwise accumulate to toxic levels within the body. For example, it has been estimated that the time required to excrete one-half of a single dose of benzene would be about 100 years if the liver did not metabolize it. The primary disadvantage of the liver's role as the main organ metabolizing chemicals, however, is that toxic reactive chemicals or short-lived intermediates can be formed during the biotransformation process. As a result the liver is frequently the major target of the toxicity associated with the reactive or toxic chemical metabolites it generates while attempting to convert a chemical into a more water-soluble and more easily eliminated form.

Morphological Consideration

The liver can be described as a large mass of cells packed around vascular trees of arteries and veins (see Figure 6.2). Blood supply to the liver comes from the hepatic artery and the portal vein, the former normally supplying about 20% of blood reaching the liver and the latter about 80%. Terminal branches of the hepatic artery and portal vein are found together with the bile duct (Figure 6.2). In cross section, these three vessels are called the *portal triad*. Blood is collected in the terminal hepatic venules, which drain into the hepatic vein. The functional microanatomy can be viewed in different ways. In one view, the basic unit of the liver is termed the *lobule*.

Blood enters the lobule from the hepatic artery and portal veins, traverses the lobule through hepatic sinusoids, and exits through a hepatic venule. In the typical lobule view, cells near the portal vein are termed *periportal*, while those near the hepatic venule are termed *perivenular*. The hepatic venule is visualized as occupying the center of the lobule, and cells surrounding the venule are sometimes termed *centrilobular*, while those farther away, near the portal triad, are called *peripheral lobular*. Rappaport proposed a different view of hepatic anatomy in which the basic anatomical unit is called the *simple liver acinus*. In this view (Figure 6.3, left), cells within the acinus are divided into zones. The area adjacent to small vessels radiating from the portal triad is zone 1. Cells in zone 1 are first to receive blood through the sinusoids. Blood then travels past cells in zones 2 and 3 before reaching the hepatic venule. As can be seen in Figure 6.3, zone 3 is roughly analogous to the centrilobular region of the classic lobule, since it is closest to the central vein. Zone 3 cells from adjacent acini form a star-shaped pattern around this vessel. Zone 1 cells surround the terminal afferent branches of the portal vein and hepatic artery, and are often stated as occupying the *periportal* region, while cells between zones 1 and 3 (i.e., in zone 2) are said to occupy the *midzonal* region. A modification of the typical lobule and acinar models has been provided by Lamers and colleagues (Figure 6.3, right). Based on histopathological and immunohistochemical studies, they propose that zone 3 should be viewed as a circular, rather than star-shaped, region surrounding the central vein. Zone 1 cells surround

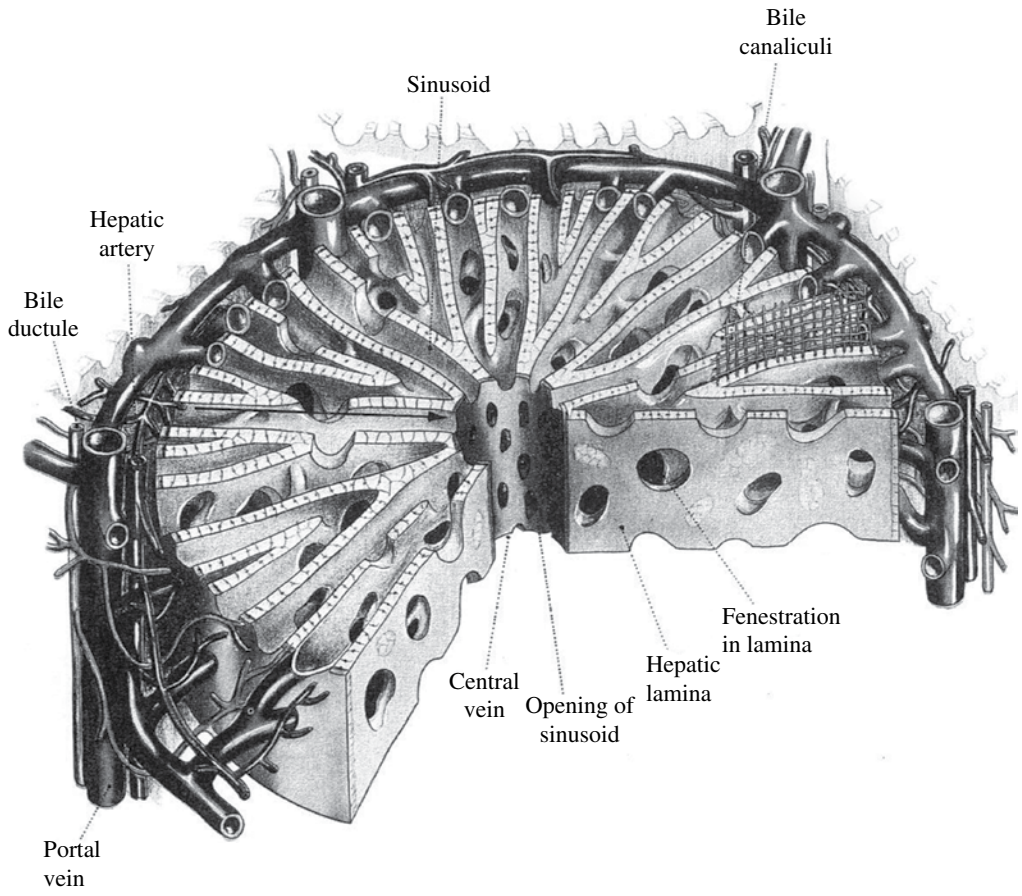


FIGURE 6.2 Hepatic architecture, showing arrangement of blood vessels and cords of liver cells. Reproduced with permission from Hamilton (1976). © 1976 Elsevier.

the portal tracts, and the zone 1 cells from adjacent acini merge to form a reticular pattern. As with the Rappaport model, cells in zone 3 may be described as centrilobular (matching closely the classic lobular terminology), cells in zone 1 as periportal, and the cells in zone 2 in between as midzonal.

Each of these viewpoints has in common a recognition that the cells closest to the arterial blood supply receive the highest concentrations of oxygen and nutrients. As blood traverses the lobule, concentrations of oxygen and nutrients diminish. Differences in oxygen tension and nutrient levels are reflected in differing morphology and enzymatic content between cells in zones 1 and 3. Consistent with their greater access to oxygen, hepatocytes in zone 1 are better adapted to aerobic metabolism. They have greater respiratory activity, greater amino acid utilization, and higher levels of fatty acid oxidation. Glucose formation from gluconeogenesis or from the breakdown of glycogen predominates in zone 1 cells, and the maximum secretion of bile acids occurs here. On the other hand, most forms of the biotransformation enzyme cytochrome P450 are found in highest concentrations in zone 3 cells. As the site of biotransformation for most drugs and chemicals, zone 3 cells have the greatest responsibility for their detoxification. This also means that zone 3 cells are often the primary targets for chemicals that are bioactivated by these enzymes to toxic metabolites in the liver.

There are several types of liver cells. *Hepatocytes*, or *parenchymal cells*, constitute approximately 75% of the total cells in the human liver. They are relatively large cells and make up the bulk of the hepatic lobule. By virtue of their numbers and their extensive xenobiotic metabolizing activity, these cells are the principal targets for hepatotoxic chemicals. The sinusoids are lined with endothelial cells. These cells are small but numerous, making up most of the remaining cells in the liver. The hepatic microvasculature also contains resident macrophages, called *Kupffer cells*. Although comparatively few in number, these cells play an important role in phagocytizing microorganisms and foreign particulates in the blood. While these cells are a part of the liver, they are also part of the immune system. They are capable of releasing reactive oxygen species and cytokines, and play an important role in inflammatory responses in the liver. The liver also contains *Ito cells* (also termed *fat-storing cells*, *parasinusoidal cells*, or *stellate cells*), which lie

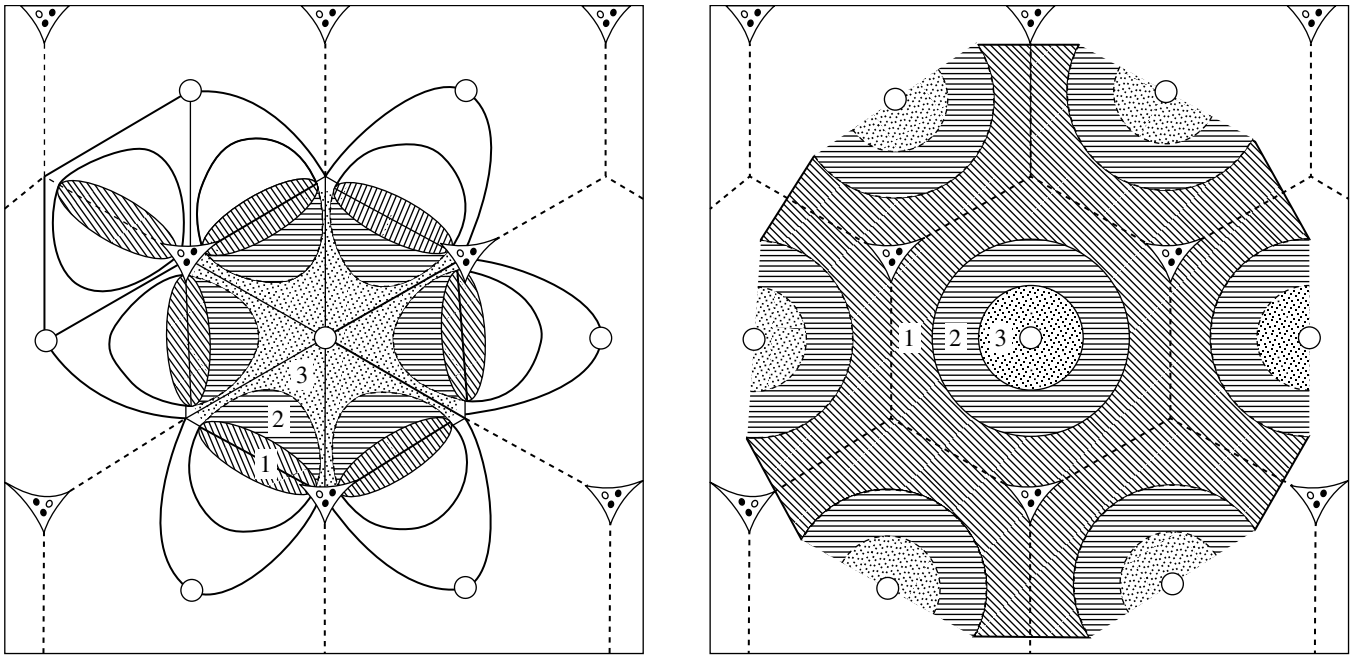


FIGURE 6.3 Alternative views of the liver acinus. Reproduced with permission from Lamers et al. (1989). © 1989 American Association for the Study of Liver Diseases.

between parenchymal and endothelial cells. These cells appear to be important in producing collagen and in vitamin A storage and metabolism.

6.2 TYPES OF LIVER INJURY

All chemicals do not produce the same type of liver injury. Rather, the type of lesion or effect observed is dependent on the chemical involved, the dose, and the duration of exposure. Some types of injury are the result of acute toxicity to the liver, while others appear only after chronic exposure or treatment. Basic types of liver injury include the anomalies described in the following paragraphs.

Hepatocellular Degeneration and Death

Many hepatotoxicants are capable of injuring liver cells directly, leading to cellular degeneration and death. A variety of organelles and structures within the liver cell can be affected by chemicals. Principal targets include the following:

1. *Mitochondria.* These organelles are important for energy metabolism and synthesis of ATP. They also accumulate and release calcium, and play an important role in calcium homeostasis within the cell. When mitochondria become damaged, they often lose the ability to regulate solute and water balance, and undergo swelling that can be observed microscopically. Mitochondrial membranes can become distorted or rupture, and the density of the mitochondrial matrix is altered. Mitochondrial dysfunction is one of the most important causes of drug-induced hepatotoxicity. Examples of chemicals that show damage to hepatic mitochondria include carbon tetrachloride, cocaine, dichloroethylene, ethionine, hydrazine, and phosphorus.
2. *Plasma Membrane.* The plasma membrane surrounds the hepatocyte and is critically important in maintaining the ion balance between the cytoplasm and the external environment. This ion balance can be disrupted by damage to plasma membrane ion pumps, or by loss of membrane integrity causing ions to leak in or out of the cell following their concentration gradients. Loss of ionic control can cause a net movement of water into the cell, resulting in cell swelling. Blisters or "blebs" in the plasma membrane may also occur in response to chemical toxicants. Examples of chemicals that show damage to plasma membrane include acetaminophen, ethanol, mercury, and phalloidin.
3. *Endoplasmic Reticulum.* The endoplasmic reticulum is responsible for the synthesis of proteins and phospholipids in the hepatocyte. It is the principal site of biotransformation of foreign chemicals and, along with the mitochondria, sequesters and releases calcium ions to promote calcium homeostasis. As discussed in Chapter 2, hepatic biotransformation enzyme activity is substantially increased in response to treatment or exposure to a variety of chemicals. Many of these enzymes, including cytochrome P450, are located in the

endoplasmic reticulum, which undergoes proliferation as part of the enzyme induction process. Because the endoplasmic reticulum is the site within the cell of most oxidative metabolism of foreign (xenobiotic) chemicals, it is also the site where reactive metabolites from these chemicals are formed. This makes it a logical target for toxicity for chemicals that produce injury through this mechanism. Morphologically, damage to the endoplasmic reticulum often appears in the form of dilation. Examples of chemicals that show damage to the endoplasmic reticulum include acetaminophen, bromobenzene, carbon tetrachloride, and cocaine.

4. *Nucleus.* There are several ways in which the nuclei can be damaged by chemical toxicants. Some chemicals or their metabolites can bind to DNA, producing mutations (see Chapter 14). These mutations can alter critical functions of the cell leading to cell death, or can contribute to malignant transformation of the cell to produce cancer. Some chemicals appear to cause activation of endonucleases, enzymes located in the nucleus that digest chromatin material. This leads to uncontrolled digestion of the cell's DNA—obviously not conducive to normal cell functioning. Some chemicals cause disarrangement of chromatin material within the nucleus. Morphologically, damage to the nucleus appears as alterations in the nuclear envelope, in chromatin structure, and in the arrangement of nucleoli. Examples of chemicals that produce nuclear alterations include aflatoxin B, beryllium, ethionine, galactosamine, and nitrosamines.
5. *Lysosomes.* These subcellular structures contain digestive enzymes (e.g., proteases) and are important in degrading damaged or aging cellular constituents. In hepatocytes injured by chemical toxicants, their numbers and size are often increased. Typically, this is not because they are a direct target for chemical attack, but rather they reflect the response of the cell to the need to remove increased levels of damaged cellular materials caused by the chemical.

Not all hepatocellular toxicity leads to cell death. Cells may display a variety of morphological abnormalities in response to chemical insult and still recover. These include cell swelling, dilated endoplasmic reticulum, condensed mitochondria and chromatin material in the nucleus, and blebs on the plasma membrane. More severe morphological changes are indicative that the cell will not recover, and will proceed to cell death, that is, undergo *necrosis.* Examples of morphological signs of necrosis are massive swelling of the cell, marked clumping of nuclear chromatin, extreme swelling of mitochondria, breaks in the plasma membrane, and the formation of cell fragments.

Necrosis from hepatotoxic chemicals can occur within distinct zones in the liver, be distributed diffusely, or occur massively. Many chemicals produce a zonal necrosis; that is, necrosis is confined to a specific zone of the hepatic acinus. Table 6.1 provides examples of drugs and chemicals that produce hepatic necrosis and the characteristic zone in which the lesion occurs. Figure 6.4 shows an example of zone 3 hepatic necrosis from acetaminophen. Confinement of the lesion to a specific zone is thought to be a consequence of the mechanism of toxicity of these agents and the balance of activating

TABLE 6.1 Drugs and Chemicals That Produce Zonal Hepatic Necrosis

	Site of Necrosis		
Chemical	Zone 1	Zone 2	Zone 3
Acetaminophen			X
Aflatoxin	X		X
Allyl alcohol	X		
Alloxan	X		
α-Amanitin			X
Arsenic, inorganic			X
Beryllium		X	
Botulinum toxin			X
Bromobenzene			X
Bromotrichoromethane			X
Carbon tetrachloride			X
Chlorobenzenes			X
Chloroform			X
Chloroprene			X
Cocaine[a]		X	X
Dichlorpropane			X
Dioxane			X
DDT			X
Dimethylnitrosamine			X
Dinitrobenzene			X
Dinitrotoluene			X
Divinyl ether			X
Ethylene dibromide			X
Ethylene dichloride			X
Ferrous sulfate	X		
Fluoroacetate			X
Iodobenzene			X
Iodoform			X
Manganese compounds	X		
Methylchloroform			X
Naphthalene			X
Ngaione		X	
Paraquat		X	
Phalloidin			X
Pyridine			X
Pyrrolidizine alkaloids			X
Rubratoxin			X
Tannic acid			X
Thioacetamide			X
Urethane			X
Xylidine			X

Source: Adapted from Cullen and Reubner (1991).

[a]Necrosis is shifted to zone 1 in phenobarbital-pretreated animals.

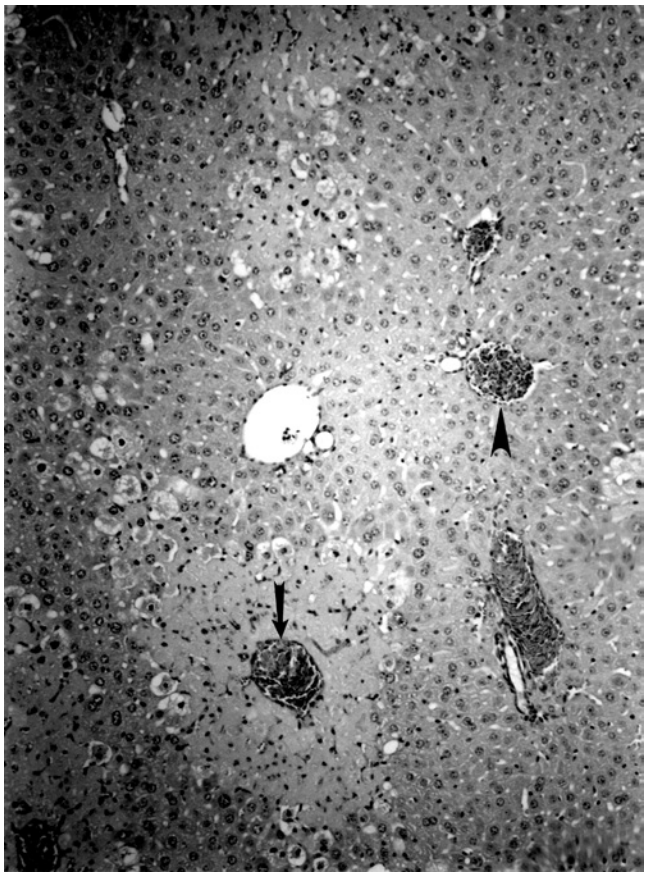


FIGURE 6.4 Liver section from mouse given a hepatotoxic dose of acetaminophen. With acetaminophen, liver cell swelling and death characteristically occur in regions around the central vein (zone 3, arrow). Cells near the portal triad (arrow head) are spared.

and inactivating enzymes or cofactors. Interestingly, there are a few chemicals for which the zone of necrosis can be altered by treatment with other chemicals. These include cocaine, which normally produces hepatic necrosis in zone 2 or 3 in mice, but in phenobarbital-pretreated animals causes necrosis in zone 1. Limited observations of liver sections from humans experiencing cocaine hepatotoxicity are consistent with this shift produced by barbiturates. The reason for the change in site of necrosis with these chemicals is unknown.

Necrotic cells produced by some chemicals are distributed diffusely throughout the liver, rather than being localized in acinar zones. Galactosamine and the drug methylphenidate are examples of chemicals that produce a diffuse necrosis. Diffuse necrosis is also seen in viral hepatitis and some forms of idiosyncratic liver injury. The extent of necrosis can vary considerably. When most of the cells of the liver are involved, this is termed *massive necrosis.* As the name implies, this involves destruction of most or all of the hepatic acinus. Not all the acini in the liver are necessarily affected to the same extent, but at least some acini will have necrosis that extends across the lobule from the portal triad to the hepatic vein, called *bridging necrosis*. Massive necrosis is not so much a characteristic of specific hepatotoxic chemicals as of their dose.

Because of the remarkable ability of the liver to regenerate itself, it is able to withstand moderate zonal or diffuse necrosis. Over a period of several days, necrotic cells are removed and replaced with new cells, restoring normal hepatic architecture and function. If the number of damaged cells is too great, however, the liver's capacity to restore itself becomes overwhelmed, leading to hepatic failure and death.

Another form of cell death is *apoptosis*, or programmed cell death. Apoptosis is a normal physiological process used by the body to remove cells when they are no longer needed or have become functionally abnormal. In apoptosis, the cell "commits suicide" through activation of its endonucleases, destroying the cell's DNA. Apoptotic cells are morphologically distinct from cells undergoing necrosis as described earlier. Unlike cells undergoing necrosis, which swell and release their cellular contents, apoptotic cells generally retain plasma membrane integrity and shrink, resulting in condensed cytoplasm and dense chromatin in the nucleus. Apoptosis is often a slower process than necrosis. There are normally few apoptotic cells in liver, but the number may be increased in response to some hepatotoxic chemicals, notably thioacetamine and ethanol. Also, some chemicals produce hypertrophy, or growth of the liver beyond its normal size. Examples include lead nitrate and phenobarbital. When exposure or treatment with these agents has ended, the liver will return to its normal size. During this phase, the number of apoptotic cells is increased, reflecting an effort by the liver to reduce its size, in part by eliminating some of its cells.

Hepatocellular degeneration can lead to autophagy. Autophagy is a process through which damaged proteins and organelles are engulfed by autophagosomes within the cell, which fuse with lysosomes. Within these vesicles, the proteins and organelles are broken down into component amino acids, lipids, and nucleotides that can then be recycled by the cell. Although a normal part of cell physiology, high levels of autophagy can lead to cell death. The molecular pathways that control autophagy and apoptosis are different, as is the appearance of cells undergoing death from these two processes. In apoptosis, the cytoskeleton is lost early while organelles are not affected until late in the process. The opposite occurs in autophagy, where organelles are lost early as autophagic vesicles accumulate, while the cytoskeleton is not affected until later.

Drugs and chemicals can produce hepatocellular degeneration and death by many possible mechanisms. For some hepatotoxicants, the mechanism of toxicity is reasonably well established. For example, galactosamine is thought to cause cell death by depleting uridine triphosphate, which is

essential for synthesis of membrane glycoproteins. For most hepatotoxicants, however, key biochemical effects responsible for hepatocellular necrosis remain uncertain. The search for a broadly applicable mechanism of hepatotoxicity has yielded several candidates:

Lipid Peroxidation Many hepatotoxicants generate free radicals in the liver. In some cases, such as carbon tetrachloride, the free radicals are breakdown products of the chemical generated by its cytochrome P450–mediated metabolism in the liver. In other cases, the chemical causes a disruption in oxidative metabolism within the cell, leading to the generation of reactive oxygen species. An important potential consequence of free radical formation is the occurrence of lipid peroxidation in membranes within the cell. Lipid peroxidation occurs when free radicals attack the unsaturated bonds of fatty acids, particularly those in phospholipids. The free radical reacts with the fatty acid carbon chain, abstracting a hydrogen. This causes a fatty acid carbon to become a radical, with rearrangement of double bonds in the fatty acid carbon chain. This carbon radical in the fatty acid reacts with oxygen in a series of steps to produce a lipid hydroperoxide and a lipid radical that can then react with another fatty acid carbon. The peroxidation of the lipid becomes a chain reaction, resulting in fragmentation and destruction of the lipid. Because of the importance of phospholipids in membrane structure, the principal consequence of lipid peroxidation for the cell is loss of membrane function. The reactive products generated by lipid peroxidation can interact with other components of the cell as well, and this could also contribute to toxicity.

The list of chemicals that produce lipid peroxidation as part of their hepatotoxic effects is extensive, and includes halogenated hydrocarbons (e.g., carbon tetrachloride, chloroform, bromobenzene, tetrachloroethene), alcohols (e.g., ethanol, isopropanol), hydroperoxides (e.g., *tert*-butylhydroperoxide), herbicides (e.g., paraquat), and a variety of other compounds (e.g., acrylonitrile, cadmium, cocaine, iodoacetamide, chloroacetamide, sodium vanadate). Consequently, it is an attractive common mechanism of hepatotoxicity. There is some question, however, as to whether it is the most important mechanism of toxicity for these chemicals. For some of these hepatotoxic compounds, experiments have been conducted in which lipid peroxidation was blocked by concomitant treatment with an antioxidant. In many cases, hepatotoxicity still occurred. This argues that for at least some agents, lipid peroxidation may contribute to their hepatotoxicity, but is not sufficient to explain all of their toxic effects on the liver.

Irreversible Binding to Macromolecules Most of the conventional hepatotoxicants must be metabolized in order to produce liver toxicity, producing one or more chemically reactive metabolites. These reactive metabolites bind irreversibly to cellular macromolecules—primarily proteins, but in some cases also lipids and DNA. This binding precedes most manifestations of toxicity, and the extent of binding often correlates well with toxicity. In fact, histopathology studies with some of these chemicals have found that only cells with detectable reactive metabolite binding undergo necrosis. Examples of hepatotoxic chemicals that produce reactive metabolites include acetaminophen, bromobenzene, carbon tetrachloride, chloroform, cocaine, and trichloroethylene.

It is certainly plausible that irreversible binding of a toxicant to a critical protein or other macromolecule in the cell could lead to loss of its function, and the fact that binding precedes most, if not all, toxic responses in the cell makes it a logical initiating event. However, demonstrating precisely how irreversible binding causes cell death has been extremely challenging. Several studies have been conducted attempting to identify the macromolecular targets for binding and to determine whether this binding results in an effect that could lead to cell death. Acetaminophen, in particular, has been studied in this regard. While several proteins bound by the acetaminophen reactive metabolite, *N*-acetyl-*p*-benzoquinone imine, have been identified, none as yet has been clearly shown to be instrumental in acetaminophen-induced hepatic necrosis. Without identification of the critical target(s) for irreversible binding for hepatotoxicants, this remains an attractive but unproven mechanism.

Loss of Calcium Homeostasis Intracellular calcium is important in regulating a variety of critical intracellular processes, and the concentration of calcium within the cell is normally tightly regulated. The plasma membrane actively extrudes calcium ion from the cell to maintain cytosolic concentrations at a low level compared with the external environment (the ratio of intracellular to extracellular concentration is about 1:10,000). Both the mitochondria and endoplasmic reticulum are capable of sequestering and releasing calcium ion as needed to modulate calcium concentrations for normal cell functioning. Loss of control of intracellular calcium can lead to a sustained rise in intracellular calcium levels, which, in turn, disrupts mitochondrial metabolism and ATP synthesis, damages microfilaments used to support cell structure, and activates degradative enzymes within the cell. These events could easily account for cell death from hepatotoxic chemicals.

Early studies of toxic effects of chemicals on liver cells in culture suggested that an influx of calcium from outside the cell (e.g., from plasma membrane failure) was responsible for their toxic effects. Later experiments showed that this was probably not the case, but, nonetheless, supported disregulation of intracellular calcium as a key event in toxicity. Intracellular calcium levels were observed to rise substantially in response to a number of hepatotoxicants,

apparently due to chemical effects on mitochondria and/or the endoplasmic reticulum leading to loss of control of intracellular calcium stores. Impaired extrusion of calcium out of the cell by the plasma membrane might also be important, at least for some chemicals. In general, increases in intracellular calcium preceded losses of viability, suggesting a cause–effect relationship. It is sometimes difficult, however, to discern to what extent elevated calcium levels are the cause of, or merely the result of, cytotoxicity.

Immune Reactions This mechanism of hepatotoxicity is not common, but is nonetheless important. Characteristically, an initial exposure is required that does not produce significant hepatotoxicity—a sensitizing event. Subsequent exposure to the drug or chemical can lead to profound liver toxicity that may be accompanied by hepatic inflammation. Consistent with a hypersensitivity reaction, there is little evidence of a dose–response relationship, and even small doses can trigger a reaction. This response is usually rare and difficult to predict; hence it is often considered an idiosyncratic reaction. Typically, this kind of hepatotoxicity for a drug or chemical is very difficult to demonstrate in laboratory animals, and unfortunately becomes known only after widespread use or exposure in humans.

Perhaps the most familiar example of a drug or chemical producing this type of hepatoxicity is the general anesthetic halothane. Studies suggest that halothane is metabolized to a reactive metabolite that binds with proteins. These proteins become expressed on the cell surface where they are recognized by the immune system as being foreign. The immune system then mounts a cell-mediated response, resulting in destruction of the hepatocytes. This response, called *halothane hepatitis*, seldom occurs (only about 1 in 10,000 anesthetic administrations in adults) but has a 50% mortality rate. A similar phenomenon has been observed with other drugs, including diclofenac.

Fatty Liver

Many chemicals produce an accumulation of lipids in the liver, called *fatty liver or steatosis*. Examples of chemicals that produce fatty liver are provided in Table 6.2. Just as hepatocellular necrosis preferentially occurs in specific acinar zones in response to certain chemicals, so does fatty liver. For example, zone 1 is the primary site of lipid accumulation from white phosphorus, while zone 3 is where most of the lipid accumulation is observed with tetracycline and ethanol. The lipid accumulates in vacuoles within the cytoplasm, and these vacuoles are usually present as either one large, clear vacuole (called *macrovesicular steatosis*) or numerous small vacuoles (*microvesicular steatosis*). The type of steatosis (macro- or microvesicular) is characteristic of specific hepatotoxicants and, in some cases, of certain diseases or conditions. For example, microvesicular steatosis has been associated with tetracycline, valproic acid, salicylates, aflatoxin, dimethylformamide, and some of the antiviral nucleoside analogs used to treat HIV. It is also associated with Reye's syndrome and fatty liver of pregnancy. Macrovesicular steatosis has been associated with antimony, barium salts, carbon disulfide, dichloroethylene, ethanol, hydrazine, methyl and ethyl bromide, thallium, and uranium compounds.

TABLE 6.2 Drugs and Chemicals That Produce Fatty Liver

Antimony	Ethyl chloride
Barium salts	Hydrazine
Borates	Methyl bromide
Carbon disulfide	Orotic acid
Chromates	Puromycin
Dichloroethylene	Safrole
Dimethylhydrazine	Tetracycline
Ethanol	Thallium compounds
Ethionine	Uranium compounds
Ethyl bromide	White phosphorus

There are several potential chemical effects that can give rise to accumulation of lipids in the cell.

These include:

1. *Inhibition of Lipoprotein Synthesis.* A number of chemicals are capable of inhibiting synthesis of the protein moiety needed for synthesis of lipoproteins in the liver. These include carbon tetrachloride, ethionine, and puromycin.
2. *Decreased Conjugation of Triglycerides with Lipoproteins.* Another critical step in lipoprotein synthesis is conjugation of the protein moiety with triglyceride. Carbon tetrachloride, for example, can interfere with this step.
3. *Interference with Very-Low-Density Lipoprotein (VLDL) Transfer.* Inhibition of transfer of VLDL out of the cell results in its accumulation. Tetracycline is an example of an agent that interferes with this transfer.
4. *Impaired Oxidation of Lipids by Mitochondria.* Oxidation of nonesterified fatty acids is an important aspect of their hepatocellular metabolism, and decreased oxidation can contribute to their accumulation within the cell. Carbon tetrachloride, ethionine, and white phosphorus have been shown to inhibit this oxidation.
5. *Increased Synthesis of Fatty Acids.* The liver is capable of synthesizing fatty acids from acetyl-CoA (coenzyme A), and increased fatty acid synthesis can increase the lipid burden of the cells. Ethanol is an example of a chemical that produces this effect.

Other possible mechanisms might contribute to fatty liver, such as increased uptake of lipids from the blood by the liver, but the role of these processes in drug- or chemical-induced steatosis is less clear. The mechanisms listed here are not mutually exclusive. Indeed, it is likely that many of the chemicals that produce steatosis do so by producing more than one of these effects.

Fatty liver may occur by itself, or in conjunction with hepatocellular necrosis. Many chemicals produce a lesion that consists of both effects. Examples include aflatoxins, amanitin, arsenic compounds, bromobenzene, carbon tetrachloride, chloroform, dimethylnitrosamine, dinitrotoluene, DDT, dichloropropane, naphthalene, pyrrolizidine alkaloids, and tetrachloroethane. Drug- or chemical-induced steatosis is reversible when exposure to the agent is stopped.

Phospholipidosis is a special form of steatosis. It results from accumulation of phospholipids in the hepatocyte, and can be caused by some drugs as well as by inborn errors in phospholipid metabolism. Liver sections from patients with phospholipidosis reveal enlarged hepatocytes with "foamy" cytoplasm. Often this condition progresses to cirrhosis. Examples of drugs associated with phospholipidosis include amiodarone, chlorphentermine, and 4,4′-diethylaminoethoxyhexoestrol.

Cholestasis

The term *cholestasis* refers to decreased or arrested bile flow. Many drugs and chemicals are able to produce cholestatic injury, and examples are listed in Table 6.3. There are several potential causes of impaired bile flow, many of which can become the basis for drug- or chemical-induced cholestasis. Some of these are related to loss of integrity of the canalicular system that collects bile and carries it to the gall bladder, while others are related to the formation and secretion of bile. For example, α-naphthylisothiocyanate disrupts the tight junctions between hepatocytes that help form the canaliculi, the smallest vessels of the bile collection system. This causes a leakage of bile contents out of the canaliculi into the sinusoids. Other toxicants, such as methylene dianiline and paraquat, impede bile flow by damaging the bile ducts. The primary driving force for bile formation is the secretion of bile acids into the canalicular lumen. This requires uptake of bile acids from the blood into hepatocytes, and then transport into the canaliculus. Anabolic steroids are an example of a class of compounds that produce cholestatic injury by inhibiting these transport processes.

Some cholestatic injury can be expected whenever there is severe hepatic injury of any type. This is because normal bile flow requires functioning hepatocytes as well as a reasonably intact cellular architecture in the liver. Whenever this is disrupted, some impairment of bile flow can be expected as a secondary consequence. Many agents produce primarily hepatic necrosis with perhaps limited cholestasis (see Table 6.1), others produce primarily cholestasis with some necrosis (chlorpromazine and erythromycin are examples), and still others are capable of producing cholestasis with little or no damage of the hepatocytes. The contraceptive and anabolic steroids are examples of this last category of agents.

TABLE 6.3 Drugs and Chemicals That Produce Cholestasis

Amitryptyline	Ethanol
Ampicillin	Haloperidol
Arsenicals, organic	Imipramine
Barbiturates	Methylene dianiline
Carbamazepine	Methyltestosterone
Chlorpromazine	α-Naphthylisothiocyanate
Cimetidine	Norandrostenolone
Cyproheptadine	Paraquat
4,4-Diaminodiphenylmethane	Phalloidin
4,4-Diaminodiphenylamine	Phenytoin
1,1-Dichloroethylene	Prochlorperazine
Dinitrophenol	Tolbutamide
Erythromycin estolate	Troleandomycin
Estrogens	

Vascular Injury

Cells lining the vasculature within the liver are also potential targets for hepatotoxicants. Injury of vascular cells leads to occlusion (impaired blood flow), which in turn leads to hypoxia. Cells in zone 3 are most vulnerable, since the oxygenation of blood reaching these cells is low even under normal conditions. Typically, hypoxia results in necrosis, and continuing injury over time leads to fibrosis. Severe cases can result in fatal congestive cirrhosis. There are several examples of chemicals known to produce hepatic *venoocclusive disease* including many of natural origin such as pyrrolizidine alkaloids in herbal teas. Oral contraceptives and some anticancer drugs have also been associated with this effect.

Peliosis hepatis is another vascular lesion characterized by the presence of large, blood-filled cavities. It is unclear why these cavities form, but there is reason to suspect that it may be due to a weakening of sinusoidal supporting membranes. Use of anabolic steroids has been associated with this effect. Although patients with peliosis hepatis are usually without symptoms, the cavities occasionally rupture causing bleeding into the abdominal cavity.

Cirrhosis

Chronic liver injury often results in the accumulation of collagen fibers within the liver, leading to fibrosis. Fibrotic tissue accumulates with repeated hepatic insult, making it difficult for the liver to replace damaged cells and still maintain normal hepatic architecture. Fibrous tissue begins

to form walls separating cells. Distortions in hepatic microcirculation causes cells to become hypoxic and die, leading to more fibrotic scar tissue. Ultimately, the organization of the liver is reduced to nodules of regenerating hepatocytes surrounded by walls of fibrous tissue. This condition is called *cirrhosis*. Hepatic cirrhosis is irreversible and carries with it substantial medical risks. Blood flow through the liver becomes obstructed, leading to portal hypertension. To relive this pressure, blood is diverted past the liver through various shunts not well suited for this purpose. It is common for vessels associated with these shunts to rupture, leading to internal hemorrhage. Even without hemorrhagic episodes, the liver may continue to decline until hepatic failure occurs.

The ability of chronic ethanol ingestion to produce cirrhosis is widely appreciated. Occupational exposures to carbon tetrachloride, trinitrotoluene, tetrachloroethane, and dimethylnitrosamine have also been implicated as causing cirrhosis, as well as the medical use of arsenicals and methotrexate. Some drugs (e.g., methyldopa, nitrofurantoin, isoniazid, diclofenac) produce an idiosyncratic reaction resembling viral hepatitis. This condition, termed *chronic active hepatitis*, can also lead to cirrhosis if the drug is not withdrawn.

Tumors

Many chemicals are capable of producing tumors in the liver, particularly in laboratory rodents. In fact, in cancer rodent bioassays for carcinogenicity, the liver is the most common site of tumorigenicity. Hepatic tumors may be benign or malignant. Conceptually, the distinction between them is that benign tumors are well circumscribed and do not metastasize (i.e., do not invade other tissues). Malignant tumors, on the other hand, are poorly circumscribed and are highly invasive (see Chapter 15 for additional discussion on benign and malignant tumors). Benign tumors, despite their name, are capable of producing morbidity and mortality. However, they are easier to manage and have a much better prognosis than malignant tumors.

Tumors are also classified by the tissue of origin, that is, whether they arise from epithelial or mesenchymal tissue, and by the specific cell type from which they originate. The nomenclature for naming tumors is complex, and the reader is referred elsewhere for a complete discussion of the topic. Basically, malignant tumors arising from epithelial tissue are termed *carcinomas*, while malignant tumors of mesenchymal origin are *sarcomas.* Thus, malignant tumors derived from hepatocytes, which are of epithelial origin, are termed *hepatocellular carcinomas*. Malignant tumors from bile duct cells, also of epithelial origin, are termed *cholangiocarcinomas* (the prefix *cholangio-* refers to the bile ducts). Cells of the vascular lining are of mesenchymal origin. Consequently, a malignant tumor in the liver arising from these cells may be called *hemangiosarcoma*. Benign tumors are also named on the basis of tissue of origin and their appearance. For example, benign tumors of epithelial origin with gland or gland-like structures are called *adenomas*, and in the liver these can occur among hepatocytes or bile duct cells. Benign tumors of fibrotic cell origin are termed *fibromas*, and those in the bile ducts are called *cholangiofibromas.*

To make things more complicated, cells go through a series of morphological changes as they progress to become a benign or malignant tumor. Thus, groups of cells that represent proliferation of liver tissue, but are not (or not yet) tumors, may be described as nodular hyperplasia, focal hepatocellular hyperplasia, or foci of hepatocellular alteration, depending on their morphological characteristics. The foci of hepatocellular alteration represent the earliest stages that can be detected microscopically. These foci are small groups of cells that are abnormal, but have no distinct boundary separating them from adjacent cells. Their growth rate is such that they produce little or no compression of surrounding cells. The abnormalities are subtle at this stage, and special stains and markers are sometimes used to help visualize them. Nodular hyperplasia is more readily observed; the group of cells is more circumscribed and compression of adjacent cells is apparent. These cells are thought to represent an intermediate step in tumor development. The significance of these lesions is not that they are associated with any clinical signs or symptoms of disease, but rather that they may represent an area from which a tumor may develop. Consequently, their appearance is important in the assessment of the ability of a drug or chemical to cause cancer. For most chemicals, only a very small percentage—or perhaps none—of the neoplastic areas will go on to produce a malignant tumor. Consequently, the issue of how to use data regarding the appearance of these lesions in the assessment of carcinogenicity of a chemical is one of considerable discussion and debate among toxicologists.

Liver tumors from chemical exposure can arise through numerous mechanisms. Some hepatocarcinogens form DNA adducts leading to mutations. Nitrosoureas and nitrosamines are examples of hepatocarcinogens thought to produce tumors through this mechanism (see also Chapters 14 and 15 for further discussion of genotoxicity and carcinogenicity). Many chemicals that produce liver tumors are not genotoxic, however, and appear to work through epigenetic mechanisms. Nongenotoxic hepatocarcinogens are many and diverse, and include tetrachlorodibenzo-*p*-dioxin, sex steroids, synthetic antioxidants, some hepatic enzyme–inducing agents (e.g., phenobarbital), and peroxisome proliferators (e.g., clofibrate). A discussion of the mechanisms underlying epigenetic carcinogenesis (e.g., inhibition of cell-to-cell communication, recurrent cellular injury, receptor interactions) is beyond the scope of this chapter, and the reader is referred to Chapter 15 for more information on this subject.

Despite the many chemicals found to produce benign and malignant liver tumors in mice and rats, relatively few

have been clearly associated with liver tumors in humans. Adenomas have been associated with the use of contraceptive steroids, and clinical and epidemiological studies implicate anabolic steroids, arsenic, and thorium dioxide as causing hepatocellular carcinoma in humans. Hemangiosarcoma is a rare tumor that has been strongly linked to occupational exposure to vinyl chloride.

6.3 EVALUATION OF LIVER INJURY

Symptoms of Liver Toxicity

As discussed earlier, liver injury may be either acute or chronic, and may involve liver cell death, hepatic vascular injury, disruption of bile formation and/or flow, or the development of benign or malignant tumors. Obviously, the signs and symptoms that accompany this array of types of liver injury can vary significantly. There are some generalizations that can be made, however. Common symptoms of liver injury include anorexia (loss of appetite), nausea, vomiting, fatigue, and abdominal tenderness. Physical examination may reveal hepatomegaly (swelling of the liver) and ascites (the accumulation of fluid in the abdominal space). Patients whose liver toxicity involves impaired biliary function may develop *jaundice*, which results from the accumulation of bilirubin in the blood and tissues. Jaundice will appear as a yellowish tint to the skin, mucous membranes, and eyes. *Pruritis*, or an itching sensation in the skin, will often accompany the jaundice.

If the injury is particularly severe, it may lead to *fulminant hepatic failure.* When the liver fails, death can occur in as little as 10 days. There are several complications associated with fulminant hepatic failure. Because the liver is no longer able to produce clotting factor proteins, albumin, or glucose, hemorrhage and hypoglycemia are common. Also, failure of the liver leads to renal failure and deterioration of the central nervous system (*hepatic encephalopathy*). Inability to sustain blood pressure and accumulation of fluid in the lungs may also result. Prognosis is poor for patients with fulminant hepatic failure, with a mortality rate of about 90%.

Morphological Evaluation

For laboratory animal studies of hepatotoxicity, histopathological examination of liver tissue by light or electron microscopy can be extremely valuable. Histopathological evaluation can provide information on the nature of the lesion and the regions of the liver affected. This, in turn, can provide insight as to the mechanism of toxicity. For example, the presence of fatty liver would suggest that the chemical may interfere with triglyceride metabolism and/or lipoprotein secretion by the liver. Hepatocellular necrosis confined to the centrilobular region might suggest bioactivation of the chemical by cytochrome P450, since most of the activity of this enzyme normally exists in centrilobular cells. Altered morphology of mitochondria as an early event in toxicity might suggest that mitochondrial toxicity is an important initiating event in the sequence of events leading up to cell death. Histopathological observations alone cannot establish the mechanism of toxicity, and additional experimentation would be required to explore the hypothesis. Nevertheless, morphological observation provides important clues, and is an integral part of any comprehensive study of potential hepatotoxicity of a chemical.

In humans, morphological evaluation of liver biopsies is sometimes used in the diagnosis and management of chronic liver toxicity, particularly liver cancer. Also, noninvasive techniques such as computerized tomography (CT) or magnetic resonance imaging (MRI) scans are used to detect liver cancer, obstructive biliary injury, cirrhosis, and venoocclusive injury to the liver.

Blood Tests

A great deal of insight into the nature and extent of hepatic injury can often be gained through tests on blood samples. There are two fundamental types of blood tests that can be performed. One type is an assessment based on measuring the functional capabilities of the liver. This can involve an evaluation of the liver's ability to carry out one or more of its basic physiological functions (e.g., glucose metabolism, synthesis of certain proteins, excretion of bilirubin) or its capacity to extract and metabolize foreign compounds from the blood. The second type of assessment involves a determination of whether there are abnormally high levels in the blood of intracellular hepatic proteins. The presence of elevated levels of these proteins in blood is presumptive evidence of liver cell destruction. Examples of these two types of tests are described here:

1. *Serum Albumin*. Albumin is synthesized in the liver and secreted into blood. Liver damage can impair the ability of the liver to synthesize albumin, and serum albumin levels may consequently decrease. The turnover time for albumin is slow, and as a result it takes a long time for impaired albumin synthesis to become evident as changes in serum albumin. For this reason, serum albumin measurements are not helpful in assessing acute hepatotoxicity. They may assist in the diagnosis of chronic liver injury, but certain other diseases can alter serum albumin levels, and the test is therefore not very specific.
2. *Prothrombin Time*. The liver is responsible for synthesis of most of the clotting factors, and a decrease in their synthesis due to liver injury results in prolonged clotting time. In terms of clinical tests, this appears as an increase in prothrombin time. Several drugs and

certain diseases also increase prothrombin time. As with serum albumin measurement, this is a relatively insensitive and nonspecific tool for detecting or diagnosing chemical-induced liver injury.

3. *Serum Bilirubin*. The liver conjugates bilirubin, a normal breakdown product of the heme from red blood cells, and secretes the glucuronide conjugate into the bile. Impairment of normal conjugation and excretion of bilirubin results in its accumulation in the blood, leading to jaundice. Serum bilirubin concentrations may be elevated from acute hepatocellular injury, cholestatic injury, or biliary obstruction but it may also be elevated in situations resulting in the hemolysis of red blood cells. This test is always included among the battery of tests to assess liver function clinically, although it is not a particularly sensitive test for acute injury.
4. *Dye Clearance Tests*. These tests involve administration of a dye that is cleared by the liver and measurement of its rate of disappearance from the blood. Delayed clearance is interpreted as evidence of liver injury. One such dye is sulfobromophthalein (Bromsulphalein or BSP). Clearance of BSP from the blood is dependent on its active transport into liver cells, conjugation with glutathione, and then active transport into the bile. Conceivably, disruption of any of these processes could result in delayed clearance, although the biliary excretion step is regarded as most critical. The test consists of administering a dose of the dye intravenously and measuring its concentration in blood spectrophotometrically over time. Another dye used for this purpose is indocyanine green (ICG). Unlike BSP, ICG is excreted into the bile without conjugation. Following an intravenous dose, the disappearance of ICG from blood can be measured with repeated blood samples or noninvasively by ear densitometry. The dye tests, although well established, are seldom used clinically.
5. *Drug Clearance Tests*. This test relies on the principle that liver injury will result in impaired biotransformation. The biotransformation capacity of the liver is assessed by following the rate of elimination of a test drug whose clearance from blood is dependent on hepatic metabolism (i.e., a drug for which other elimination processes, such as renal excretion, are insignificant). A test drug such as antipyrine, aminopyrine, or caffeine is administered, and its rate of disappearance from blood is followed over time through serial blood sampling. This rate is compared with a value considered "normal" to determine whether impaired biotransformation exists. This can also be used to test for hepatic enzyme induction, in which the rate of elimination from blood would be increased, rather than decreased as in liver injury. This test is primarily used for research purposes.
6. *Measurement of Hepatic Enzymes in Serum.* Cells undergoing acute degeneration and injury will often release intracellular proteins and other macromolecules into blood. The detection of these substances in blood above normal baseline levels signals cytotoxicity. This is true for any cell type, and for the presence of intracellular proteins in blood to be diagnostic for any particular type of cell injury (e.g., liver toxicity vs. renal toxicity vs. cardiotoxicity), the proteins must be associated rather specifically with a target organ or tissue. Fortunately, several proteins are found primarily in hepatocytes, and their presence in blood in elevated levels is the basis for some of the most commonly used tests for hepatotoxicity. Table 6.4 shows many of the most common proteins measured in these tests. The reader will note that all of these proteins are enzymes. This is not a coincidence. While any intracellular protein specific to the liver would be useful theoretically, enzymes are proteins that can be measured specifically (by measuring the rate of their particular enzyme activity) using assays that are rapid and inexpensive. In fact, the concentrations of each of these proteins are typically measured as an enzyme activity rate, rather than a true concentration per se. The aminotransferease enzymes are abundant in liver cells and are released into the bloodstream in response to chemical-induced hepatocellular injury. Alanine aminotransferase (ALT) is found in the cytosol of the liver, while two forms of aspartate aminotransferase (AST) exist, one in the cytosol and one in the mitochondria. Serum ALT and AST measurements represents two useful and relatively sensitive tests for hepatocellular injury but are not necessarily specific for toxicity to liver cells alone. The problem of not being a specific test for liver cell toxicity is seen in some of the other frequently used clinical tests for liver toxicity (see Table 6.4).

   Aminotransferase activities (ALT and AST), alkaline phosphatase activity, and gamma glutamyltransferase transpeptidase (GGTP) are included in nearly all standard clinical test suites to assess potential hepatotoxicity. The value of performing these tests is that each test may provide unique or different response patterns to different forms or causes of liver injury, and evaluating the pattern of responses can offer insight into the type of injury that has occurred. For example, severe hepatic injury from acetaminophen can result in dramatic increases in serum ALT and AST activities (up to 500 times normal values), but only modest increases in alkaline phosphatase activity. Pronounced increases in alkaline phosphatase is characteristic of cholestatic injury, where increases in ALT and AST may be limited or nonexistent. In alcoholic liver disease, AST activity is usually greater than ALT activity, but for most other forms of hepatocellular injury ALT activities are higher.

TABLE 6.4 Serum Enzyme Indicators of Hepatotoxicity

Enzyme	Acronym	Comments
Alanine aminotransferase	ALT	Found mainly in the liver; increase reflects primarily hepatocellular damage and may be a more specific indication of hepatocellular injury than ALT elevations in certain situations. It is present in other tissues in low concentrations; so, elevations of ALT are typically viewed as a specific indication of hepatocellular injury. Levels might also fluctuate during the day and in response to exercise
Aspartate aminotransferase	AST	Less specific to the liver than ALT; increase reflects primarily hepatocellular damage but it may also be elevated by nonhepatic sources like cardiac and skeletal muscle cell toxicity. In some cases these sources might need to be ruled out (e.g., in situations where AST is elevated but ALT is not)
Alkaline phosphatase	ALP	Increases reflect primarily cholestatic injury but a nonhepatic source of this enzyme is the bone, first trimester placenta, kidneys, and the intestines. So, elevated ALP levels may be seen during a nontoxic event like growth spurts in adolescents
γ-Glutamyl transferase; γ-glutamyltranspeptidase	GGTP	Increases reflect primarily cholestatic injury, although elevated in hepatocellular damage as well
5′-Nucleotidase	5′ND	Increases reflect primarily cholestatic injury
Sorbitol dehydrogenase	SDH	High specificity for liver toxicity; increase reflects primarily hepatocellular damage
Ornithine carbamoyltransferase	OCT	High specificity for liver toxicity; increase reflects primarily hepatocellular damage

Serum GGTP is an extremely sensitive indicator of hepatobiliary effects, and may be elevated simply by drinking alcoholic beverages. It is not a particularly specific indicator (it is increased by both hepatocellular and cholestatic injury) and is best utilized in combination with other tests. Increases in serum levels of ALT and AST that are higher than the corresponding changes in ALP are typically seen in hepatocellular injury. The opposite pattern is typically observed with cholestatic liver toxicity. Serum levels of enzymes such as lactate dehydrogenase have been used to evaluate liver toxicity, but this enzyme has such low specificity for the liver that interpretation of these results in a clinical setting is impossible without other confirmatory tests. It is, however, a useful indicator of liver disease in an animal research setting where the investigator knows changes in serum levels of this enzymes are specific for liver toxicity as target organs are not affected by the chemical. Other enzymes such as sorbitol dehydrogenase (SDH) and ornithine carbamoyltransferase (OCT) are quite specific to the liver.

6.4 SUMMARY

Both the anatomical location and its role as a primary site for biotransformation make the liver uniquely susceptible to drug- and chemical-induced injury. Many chemicals encountered in the workplace and the environment are capable of producing toxic effects in the liver:

- There are many types of liver injury, including hepatocellular degeneration and death (necrosis), fatty liver, cholestasis (decreased or arrested bile flow), vascular injury, cirrhosis, and tumor development.
- Hepatic injury from drugs and chemicals can arise from a variety of mechanisms. While the mechanism of toxicity for some chemicals is reasonably well established, many aspects of toxic mechanisms for most chemicals remain unclear.
- Hepatotoxic chemicals can attack a variety of subcellular targets. Principal organelles and structures affected include the plasma membrane, mitochondria, the endoplasmic reticulum, the nucleus, and lysosomes.
- Liver injury can be evaluated morphologically (microscopic examination of liver tissue) or through blood tests. Blood tests are designed to measure either the functional capacity of the liver or the appearance of intracellular hepatic contents in the blood.

REFERENCES AND SUGGESTED READING

Bjornsson ES, Jonasson JG. Drug-induced cholestasis. Clin Liver Dis 2013;17:191–209.

Boobis AR. Mode of action considerations in the quantitative assessment of tumour responses in the liver. Basic Clin Pharmacol Toxicol 2010;106:173–179.

Braubar N, Williams J. Industrial solvents and liver toxicity: risk assessment, risk factors, and mechanisms. Int J Hyg Environ Health 2002;205:479–491.

Cullen JM, Ruebner BH. A histopathologic classification of chemical-induced injury of the liver. in *Hepatotoxicity*, Meeks RG, Harrison D, Bull RJ, eds., CRC Press, Boca Raton, FL, 1991.

Davern TJ. Drug-induced liver disease. Clin Liver Dis. 2012;16:231–245.

Hamilton WJ. *Textbook of Human Anatomy*. 2nd ed. St. Louis (MO): C.V. Mosby Co.; 1976.

Ingawale DK, Madlik SK, Naik SR. Models of hepatotoxicity and the underlying cellular, biochemical and immunological mechanism(s): a critical discussion. Environ Toxicol Pharmacol 2014;37:118–133.

Jaeschke H, McGill MR, Ramachandran A. Oxidant stress, mitochondria, and cell death mechanisms in drug-induced liver injury: lessons learned from acetaminophen. Drug Metab Rev 2012;44:88–106.

Ju C, Reilly T. Role of immune reactions in drug-induced liver disease (DILI). Drug Metab Rev. 2012;44:107–115.

Lamers WH, Hilberts A, Furt E, Smith J, Jonges, GN, von Noorden CJF, Janzen JWG, Charles R, Moorman AFM. Hepatic enzyme zonation: A reevaluation of the concept of the liver acinus. Hepatology 1989;10:72–76.

Malhi H, Gores GJ, Lemasters JJ. Apoptosis and necrosis in the liver: a tale of two deaths? Hepatology 2006;43:S31–S44.

Stickel F, Kesseborhm K, Weimann R, Seitz HK. Review of liver injury associated with dietary supplements. Liver Int 2011;31:595–605.

Teschke R, Frenzel C, Glass X, Schulze J, Eickhoff A. Herbal hepatotoxicity: a critical review. Br J Clin Pharmacol 2012;75:630–636.

Wahlang B, Beier JI, Clair HB, Bellis-Jones HJ, Falkner KC, McClain CJ, Cave MC. Toxicant-induced steatohepatitis. Toxicol Pathol 2013;41:343–360.

Wallace AD. Toxic endpoints in the study of human exposure to environmental chemicals. Prog Mol Biol Transl Sci 2012;112:89–115.

7

NEPHROTOXICITY: TOXIC RESPONSES OF THE KIDNEY

Lawrence H. Lash

This chapter discusses various aspects of renal toxicology. To put the subject into its proper context, it is critical to understand the morphological, functional, and biochemical features of the kidneys that influence or largely determine the susceptibility of this organ to chemically induced toxicity. Thus, Section 7.1 summarizes some key aspects of renal structure and function and describes how these properties impact or determine the toxic responses of the kidneys to chemical or drug exposures. Section 7.2 describes the major classifications of both acute and chronic renal injury and summarizes some key aspects of how renal cells respond to these forms of injury at the cellular and molecular levels. While this chapter will by no means be a treatise on the cellular and molecular mechanisms by which nephrotoxic chemicals act, it will highlight some of the major mechanisms by which chemicals act and give an indication of which ones have been modulated to develop novel therapeutic approaches to treat nephrotoxicity. Section 7.3 summarizes common methods that are used to assess renal function and toxicity in clinical settings and in the whole animal. Section 7.4 discusses some of the most commonly used *in vitro* models from both laboratory animals and human kidney tissue that are used to screen chemicals for nephrotoxicity or investigate their mechanism of action. Finally, Section 7.5 provides some examples of nephrotoxic chemicals, and includes both environmental contaminants and therapeutically used drugs.

7.1 RENAL STRUCTURE AND PHYSIOLOGY

Overview of Kidney Structure and Function

Mammals possess two kidneys located near the lower ribs. Each kidney receives its blood supply from a single renal artery that originates in the aorta. Renal blood drains into a single renal vein that connects with the inferior vena cava. Urine normally empties from a single ureter and then into the bladder. On a gross, whole-organ level, the kidneys are subdivided into three regions: cortex, outer medulla, and inner medulla. The outer medulla is further divided into the outer stripe and inner stripe. The inner medulla is further divided into the base and tip; the latter is also called the papilla. Whereas small mammals, such as the rat, mouse, and rabbit, have a single papilla in each kidney, larger mammals, including humans, have multiple papilla per kidney. In either unipapillate or multipapillate mammals, urine exits from the tip of the papilla into the renal pelvis.

The functional unit of the kidneys is the nephron (Figure 7.1), which is a series of epithelial cells of varying types (see later) that form a tube of one-layer-thick cells enclosing a luminal compartment. Mammalian kidneys contain millions of nephrons, each of which leads into a common collecting duct. Nephrons are categorized as either superficial, midcortical, or juxtamedullary, depending on the location of their glomeruli. The renal cortex contains glomeruli, proximal convoluted and some proximal straight tubules, macula densa, cortical thick ascending limbs, distal convoluted tubules, connecting tubules, initial collecting ducts, interlobular arteries, and afferent and efferent capillary networks. The outer stripe of the outer medulla is located just below or inside the cortex and contains proximal straight tubules, medullary thick ascending limbs (mTAL), and outer medullary collecting ducts. The inner stripe of the outer medulla is located below or inside the outer stripe and contains thin descending limbs, thick ascending limbs, and outer medullary collecting ducts. The inner medulla contains thin descending limbs, thin ascending limbs, and inner medullary

Principles of Toxicology: Environmental and Industrial Applications, Third Edition. Edited by Stephen M. Roberts, Robert C. James and Phillip L. Williams.

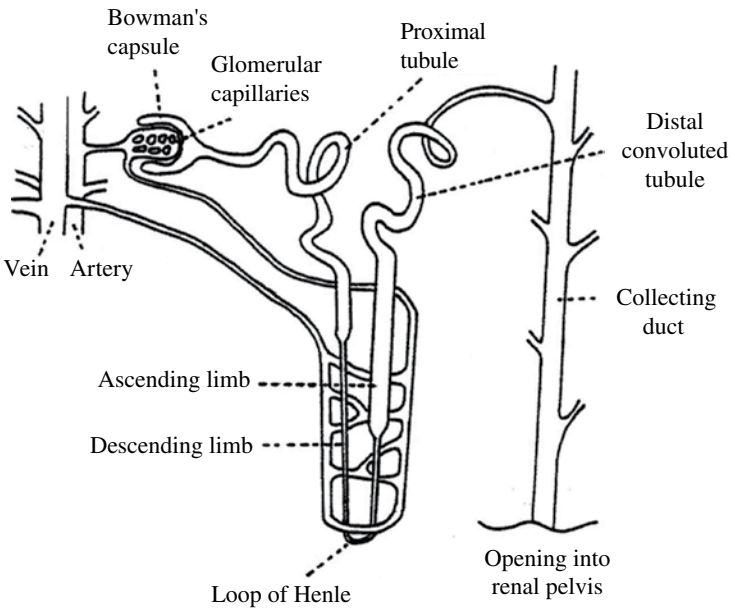


FIGURE 7.1 Schematic of the basic functional unit of the kidneys, the nephron. The scheme illustrates the basic functional unit of the kidneys, the nephron, and shows the parallel renal vasculature, the major nephron segments, and the common collecting duct into which each nephron empties.

TABLE 7.1 Filtration and Reabsorption of Water and Electrolytes by the Human Kidney

	Filtered (mEq/24 h)	Excreted (mEq/2 h)	Reabsorbed (mEq/24 h)	Reabsorbed (%)
Sodium	23,900	100	23,800	99.6
Chloride	19,800	100	19,700	99.5
Bicarbonate	5,100	2	5,098	99.9
Potassium	680	50	630	92.6
	l/24 h	**l/24 h**	**l/24 h**	
Water	169	167.5	1.5	99.1

collecting ducts. The renal vasculature surrounds the tubule and parallels the nephron.

The primary function of the kidneys is to maintain the internal environment of the body. They do this by regulating total body salt (e.g., sodium, potassium, and chloride ions), water, and acid–base balance (e.g., proton, bicarbonate, and ammonia transport and metabolism), reabsorbing nutrients (e.g., glucose, amino acids), and excreting waste products (e.g., urea, creatinine, uric acid, phosphates, sulfates). As shown in Table 7.1, the kidneys are extremely efficient at reabsorbing key electrolytes and water. The critical function of the kidneys in water conservation is exemplified by a consideration of how fluid input and output are balanced. Water input is determined largely (~88%) by ingestion via the gastrointestinal tract, with the remainder (~12%) coming from metabolism. In contrast, water output is determined primarily by the kidneys (~72%), with the remainder coming from the skin (~12%), lungs (~12%), and gastrointestinal tract (~4%).

Three general mechanisms are involved in accomplishing these key functions: glomerular filtration, tubular reabsorption, and tubular secretion. Glomerular filtration is defined as the movement of water and solutes across the glomerular capillary wall to form an ultrafiltrate of plasma. The rate at which this process occurs is called the glomerular filtration rate (GFR). A typical GFR in a normal, young, healthy adult human is 125 ml/min or 180 l/24 h. As discussed in Section 7.3, this GFR value represents a high degree of reserve function such that individuals are not considered to have compromised renal function until GFR declines to below 60 ml/min. The movement of electrolytes or metabolites from the tubular fluid back into the plasma is called tubular reabsorption, resulting in retention in the body, whereas that from plasma into the tubular fluid is

called tubular secretion, ultimately resulting in excretion into the urine.

The kidneys also possess specialized endocrine functions, such as synthesis of 1,25-dihydroxy vitamin D_3 and the protein erythropoietin. The latter is critical for synthesis of red blood cells. In fact, deficiency in the renal synthesis of erythropoietin is invariably associated with anemia. Additionally, although the liver is considered the primary organ of drug metabolism, the kidneys possess significant drug metabolism capabilities. As described in Section 7.5, some classes of chemicals are specifically bioactivated by renal enzymes to reactive metabolites that lead to their nephrotoxicity.

Nephron Structure and Heterogeneity

The mammalian kidney is a complex organ whose basic structural unit, the nephron, is composed of many cell populations, each exhibiting diverse morphological, biochemical, and functional properties (Table 7.2). Differences exist among cell types in morphology, membrane structure, energetics, and the composition of transporters and drug metabolism enzymes. As a consequence of this functional heterogeneity, exposure of the kidneys, either *in vivo* or *in vitro*, to various nephrotoxicants or to pathological conditions such as ischemia or hypoxia, produces distinct patterns of cellular injury. Thus, certain cell populations are either specific targets of or are particularly susceptible to one form of injury or another. In some cases, the specificity is due to the selective presence in a given cell population of a membrane transport system or a bioactivation or detoxication enzyme. In other cases, however, susceptibility is explained by the basic biochemical function of the cell.

For example, mTAL cells are particularly susceptible to injury from hypoxia or ischemia–reperfusion because of their relatively poor oxygenation and primary reliance on glycolysis for generation of ATP. The proximal tubules are the most common targets for chemically induced injury because they are the first nephron cell type with which filtered chemicals come in contact, they possess a large array of membrane transporters that mediate uptake and intracellular accumulation of chemicals (see later), and they contain relatively high activities of drug metabolism enzymes that can metabolize chemicals to reactive and toxic species.

Countercurrent Flow and Urinary Concentration Mechanism

The ability to generate the high osmolarity of the medullary interstitium is dependent on two principal factors: (1) active chloride or salt reabsorption in the ascending limb, and (2) the anatomical relationship between the descending and ascending limbs of the loop of Henle; the fact that both limbs are close together and that flow in each limb is in the opposite direction is a critical feature. This results in what is called *countercurrent flow*. Additionally, as noted earlier, there is a parallel *vascular structure* to the nephron structure that forms a countercurrent exchange of blood. This arrangement prevents solutes from coming out in the descending limb and being washed away but also contributes to the ability of the renal epithelium to concentrate certain chemicals to levels several-fold higher than those in either the plasma or glomerular filtrate.

Membrane Transport Processes

A critical feature of all transporting epithelia is the presence of an array of carrier proteins selectively localized on either the brush-border plasma membrane (BBM) or basolateral plasma membrane (BLM) that mediates the uptake and/or efflux of ions and metabolites. Because of nephron heterogeneity, most

TABLE 7.2 Differential Morphology, Biochemistry, and Physiology of Selected Nephron Segments

Property	Proximal tubule (PT)	Distal convoluted tubule (DCT)	Medullary thick ascending limb (mTAL)
Shape	Large, cuboidal	Small, cuboidal	Flattened
Intrarenal localization	Cortex/OSOM	Cortex	OM
Membrane structure	High-density BB microvilli, extensive BL invaginations	Few BB microvilli, few BL invaginations	Few BB microvilli, few BL invaginations
Mitochondrial density	High	Moderate	Moderate
Oxygenation	High	High	Low
Energetics	High OXPHOS, FA oxidation, gluconeogenesis	Glycolysis, low OXPHOS	Glycolysis, low OXPHOS
Transport function	OA^-, OC^+, glucose, amino acids; water-permeable	Na^+, K^+, H^+, HCO_3^-; water-permeable	Divalent cations, Na^+–K^+–Cl^-; low water permeability
Drug metabolism	CYP, FMO, GSH-dependent	Low	PGS

BB, brush-border; BL, basolateral; CYP, cytochrome P450; DCT, distal convoluted tubule; FA, fatty acid; FMO, flavin-containing monooxygenase; GSH, glutathione; mTAL, medullary thick ascending limb; OA^-, organic anion; OC^+, organic cation; OM, outer medulla; OSOM, outer stripe of outer medulla; OXPHOS, oxidative phosphorylation; PGS, prostaglandin synthase.

of the transport of drugs and environmental chemicals occurs in the proximal tubules as these cells contain the highest activities of carrier proteins for organic anions (OA^-) and cations and intermediary metabolites.

Three modes of transport are relevant to a consideration of the role of carrier-mediated transport in chemically induced nephrotoxicity (Figure 7.2). In the pathway for chemical A, the chemical is taken up into the cell from the lumen by transport across the BBM and is then reabsorbed into plasma by efflux across the BLM. In contrast, in the pathway for chemical B, the chemical is taken up into the cell from the plasma by transport across the BLM and is then excreted into the lumen and ultimately the urine by efflux across the BBM. For the pathway for chemical C, however, high-affinity uptake from plasma occurs across the BLM. Poor transport across the BBM, however, leads to intracellular accumulation of chemical C, thereby promoting toxicity.

The high efficiency by which renal cells, particularly those of the proximal tubule, take up chemicals from the plasma is associated with the presence of coupled, active transport systems in the plasma membranes (Figure 7.3). This coupling is achieved on the BLM by the action of the primary active transporter the ($Na^+ + K^+$)-stimulated ATPase, which catalyzes efflux of two Na^+ ions in exchange for uptake of three K^+ ions and is energized by the hydrolysis of ATP. This primary active transporter, so called because it uses high-energy ATP, generates a large, inwardly directed gradient of Na^+ ions across the BLM that can be used to facilitate uptake of anions. Such a facilitated, Na^+-coupled uptake occurs for 2-oxoglutarate ($2\text{-}OG^{2-}$), which is catalyzed by the sodium dicarboxylate 3 carrier (*NaC3; SLC13A3*). This is considered a secondary active transporter as it uses the Na^+ gradient generated by a primary active transporter to facilitate uptake of other substrates. The NaC3 carrier, in turn, generates an inwardly directed gradient of $2\text{-}OG^{2-}$, which is

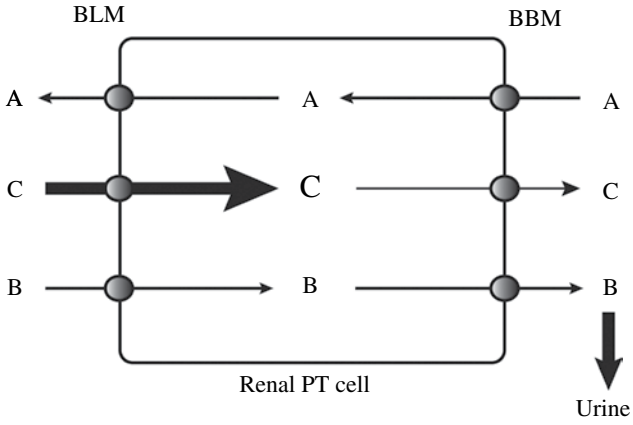


FIGURE 7.2 Modes by which plasma membrane transport determines drug disposition and toxicity in the renal proximal tubule (PT). Pathway A illustrates reabsorption of filtered chemicals by uptake into the PT cell across the brush-border plasma membrane (BBM) and then efflux from the cell into the renal plasma and interstitial space by transport across the basolateral plasma membrane (BLM). Pathway B illustrates secretion of chemicals from the renal plasma and interstitial space by successive transport across the BLM and then the BBM; the effluxed chemical is then secreted into the urine. Pathway C illustrates how chemicals in the renal plasma and interstitial space are transported into the cell across the BLM and, due to relatively poor transport out of the cell into the tubular lumen, how the chemical can accumulate in renal cells thereby enhancing toxicity.

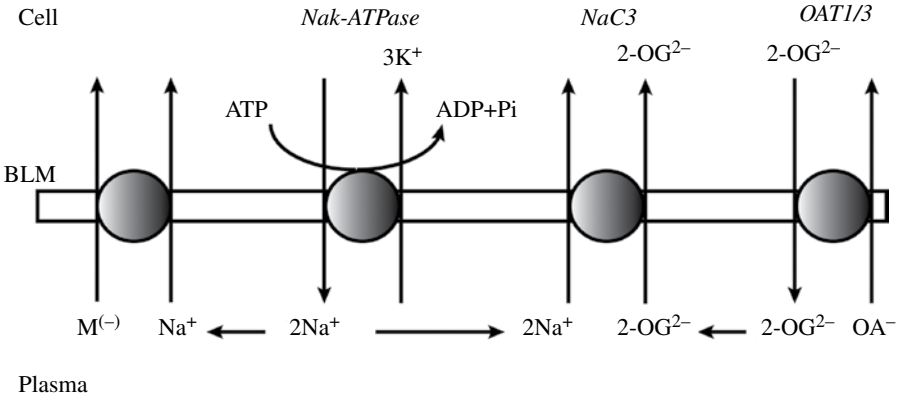


FIGURE 7.3 Coupling of primary, secondary, and tertiary active transport systems on the basolateral plasma membrane (BLM) of the renal proximal tubule. Metabolite ($M^{(-)}$) transport across the BLM is ultimately energized by direct or indirect coupling to ATP hydrolysis or coupling to ion or metabolite gradients generated by ATP hydrolysis. The ($Na^+ + K^+$)-stimulated ATPase is considered a primary active transporter as it generates Na^+ and K^+ electrochemical gradients by using the energy from ATP hydrolysis. Na^+ ions are then coupled to uptake of metabolites or 2-oxoglutarate ($2\text{-}OG^{2-}$), the latter being mediated by the sodium dicarboxylate carrier 3 (NaC3; *Slc13a3*). The NaC3 generates an electrochemical gradient of $2\text{-}OG^{2-}$, which is then coupled to the uptake of organic anions (OA^-), primarily by the OA^- transporter 1 and 3 (Oat1/3; *Slc22a6/8*). Inasmuch as Oat1 and Oat3 are energized by the $2\text{-}OG^{2-}$ gradient, which is in turn generated by the NaC3, they are termed tertiary active transporters.

used in a tertiary active transport process to catalyze uptake of OA^-, which can be mediated by, among others, the OA^- transporters 1 and 3 (*OAT1/3*; *SLC22A6/8*).

The importance of these coupled processes is highlighted by the realization that a substantial number of drugs and nephrotoxicants are OA^- and are thus transported across the BLM by OAT1 and/or OAT3. These carriers, particularly OAT1, are inhibited by the classic OA^- transport inhibitor probenecid. For many nephrotoxicants that are OA^-, their toxicity can be modulated by probenecid, demonstrating the critical role of the BLM transport step. Coadministration of multiple drugs that are all transported by these anion carriers can lead to drug–drug interactions that limit the therapeutic efficacy of these drugs and may lead to nephrotoxicity as a side effect of exposure.

Summary of Physiological Factors Contributing to the Kidneys as a Target Organ for Toxic Chemicals

The foregoing discussion highlights the various aspects of renal physiology that contribute to susceptibility of the kidneys as a target organ for many toxic chemicals. These factors are summarized in Table 7.3. As described earlier, the physiological functions of the kidneys can enhance nephrotoxicity when the chemical being handled is a nephrotoxicant. Additionally, membrane transporters and drug metabolism enzymes that are similar to those found in the liver, which is the organ usually considered the major site of drug metabolism, also exist in the kidneys, particularly in the proximal tubules.

TABLE 7.3 Factors Influencing Susceptibility of Kidneys to Chemically Induced Nephrotoxicity

Parameter	Comments
1. Blood flow	The two kidneys comprise <1% of total body weight but receive 20–25% of cardiac output.
2. Concentrating mechanisms	Countercurrent renal blood flow minimizes loss of solutes from tubule or lumen and leads to intracellular accumulation.
3. Filtration	Glomerular filtration: Ultrafiltrate of plasma; 125 ml/min in healthy, young adults.
4. Absorption and secretion	*Absorption*: Lumen to cell to blood; *Secretion*: Blood to cell to lumen; plasma membrane transporters for organic anions and cations, inorganic ions, metabolites; active transport
5. Bioactivation	Drug metabolism enzymes generate intracellular reactive metabolites.

7.2 CLASSIFICATIONS OF RENAL INJURY

In considering the manner by which chemicals can produce kidney injury, critical aspects include the dose and exposure time. This is a general point that applies not only to exposure to nephrotoxicants, but to exposure to any toxicant and any target organ. The importance of this is that exposures to relatively high doses of chemicals as a single dose or multiple doses over short periods of time (*acute exposures*) and those to relatively low doses of chemicals over longer periods of time (*chronic exposures*) result in fundamentally different organ responses.

Acute Kidney Injury

There are four basic mechanisms by which chemicals can produce acute kidney injury:

1. *Hypoperfusion or hypofiltration*: This mechanism of acute renal injury involves reductions in renal blood flow due to either renal vasoconstriction or glomerular injury, and has been associated with acute exposures to drugs such as the fungicide amphotericin B, aminoglycoside antibiotics, cyclosporin, nonsteroidal anti-inflammatory agents (NSAIDs), and radiocontrast agents.
2. *Acute tubular necrosis*: This mechanism involves direct injury to tubular epithelial cells and includes the largest and most diverse array of chemical exposures. Examples of nephrotoxic chemicals that produce direct tubular injury include aminoglycoside antibiotics, the antifungal agent amphotericin B, the analgesic acetaminophen (APAP), β-lactam antibiotics, cisplatin (CDDP), halogenated hydrocarbons, radiocontrast agents, and heavy metals such as inorganic mercury (Hg) and cadmium (Cd).
3. *Obstruction*: Radiocontrast agents can produce tubular obstruction due to the formation of urinary protein casts. This is believed to be due to the hyperosmolar nature of these agents, producing 5- to 10-fold higher concentrations of the agents in the tubular lumen than in the plasma and hemodynamic effects (i.e., renal vasoconstriction).
4. *Tubulointerstitial fibrosis*: Certain nephrotoxicants, including β-lactam antibiotics, NSAIDs, sulfonamides, and tetracycline can cause interstitial fibrosis through immunological and inflammatory effects.

Chronic Kidney Injury

Besides chemically induced carcinogenesis, chronic exposures to nephrotoxic chemicals produce renal injury primarily by two mechanisms, chronic tubulointerstitial fibrosis and papillary necrosis. Analgesics and other NSAIDs are linked to these mechanisms. The former, like acute tubulointerstitial fibrosis, involves immunological and inflammatory effects. The latter mechanism, papillary necrosis, occurs because the patterns of renal blood flow over time lead to accumulation of drug in the renal medulla.

Responses at the Cellular Level

Exposure to a nephrotoxicant can result in a broad range of responses (Figure 7.4). The precise response will vary due to four general factors: (1) the nature of the chemical, (2) exposure dose, (3) exposure duration or pattern, and (4) other factors that may alter susceptibility to renal injury.

Factor (1) It would seem intuitive that the chemical properties of the nephrotoxicant would play a major role in determining what type of response occurs at the cellular level. Indeed, some chemicals, because of their physicochemical properties (e.g., charge, lipid solubility) or their bioactivation (metabolism) to a specific reactive species, interact with very specific molecular components of the renal cell, thus producing very specific effects. Some of the examples of nephrotoxicants that will be discussed in Section 7.5 will highlight these targeted mechanisms of action.

Factor (2) A basic tenet of all Pharmacology and Toxicology is, to quote Paracelsus, that "the dose makes the poison." Thus, whereas acute cytotoxicity tends to be a response to toxic insult from high-dose exposures, altered cellular regulation more typically occurs at lower doses of toxicant exposures (cf. Figure 7.4). It is also well known that many nephrotoxic chemicals can produce both acute tubular necrosis and apoptosis, with the former occurring at relatively high doses and the latter occurring at relatively low doses.

Factor (3) Similar to dose, duration or pattern of exposure are also central components of pharmaco- and toxicokinetics. Chemical exposures can be either acute (continuous and short-term over a period of <2 h or a single exposure), subchronic (continuous or repeated exposures over a period of <3 months), or chronic (continuous or repeated exposures over a period of >3 months). Acute exposures are typically at relatively high doses (e.g., overdose toxicity) whereas chronic exposures are typically at relatively low doses (e.g., exposures to carcinogens that are environmental contaminants).

Factor (4) Many other factors may contribute to eliciting different responses to nephrotoxicants in humans. Some of these include prior or simultaneous exposure to other drugs or toxic chemicals, the existence of a disease or pathological state, or

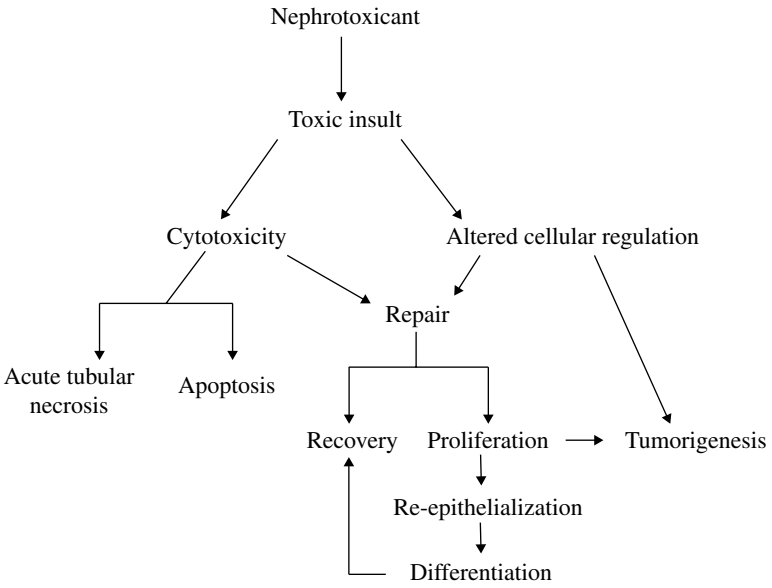


FIGURE 7.4 Generalized scheme of responses of renal cells to nephrotoxicant exposures. Upon exposure to potentially nephrotoxic chemicals or drugs, renal cells elicit a range of responses, largely dependent on the dose and duration of the exposure. Exposure to high concentrations of nephrotoxicant typically leads to cytotoxicity, which can result in acute tubular necrosis or apoptosis, although the renal proximal tubular cell can also repair sublethal injury. Exposure to relatively low doses of nephrotoxicant can produce more subtle effects, such as changes in gene expression or signal transduction pathways. This may ultimately lead to tumorigenesis or repair. The repair process typically occurs in three phases: proliferation, reepithelialization, and differentiation.

genetic polymorphisms in drug metabolism transporters or enzymes. In each case, the susceptibility of the individual is altered, thereby changing the toxic response from what might have been expected in the typical individual. In the first case, for example, prior or simultaneous exposure to other drugs or toxic chemicals may alter the response to the chemical in question because of drug–drug interactions based on such processes as competition for transport, enzyme induction or inhibition, or potentiation of toxic effect. In the second case, the exacerbation of nephrotoxicity by several diseases or pathological states has been well documented. Examples include liver disease, diabetes, and cardiovascular disease. In the third case, genetic polymorphisms may alter the ability of kidney cells to bioactivate or detoxify nephrotoxicants or to absorb or secrete them. Genetic polymorphisms, which are inherited mutations that occur in a population with a frequency of at least 1–2%, result in either altered activity of proteins or even the absence of a protein. Examples of polymorphisms that may be relevant to the ability of the kidneys, particularly the proximal tubular cells, to handle nephrotoxicants include those in drug metabolism enzymes such as cytochrome P450s (CYP) or glutathione *S*-transferases or in plasma membrane transporters such as the OA^- transporter 1 and 3 (OAT1/3; *SLC22A6/8*) or the various forms of the multidrug resistance proteins (MRPs).

In considering the cellular mechanisms by which kidney cells, in particular proximal tubular cells, respond to exposures to nephrotoxicants, it is important to understand that the renal epithelium has a significant capacity to undergo proliferation and recovery (cf. Figure 7.4). When renal epithelial cells undergo repair and proliferation, the newly generated cells are dedifferentiated, meaning that they do not express most of the proteins whose function are characteristic of mature kidney cells. There are two subsequent phases of recovery: Phase I is re-epithelialization, in which gaps in the epithelial monolayer that were caused by cell death are filled in by the new dedifferentiated cells; this is followed by differentiation of these newly generated cells. As noted in Figure 7.4, the altered cellular regulation that arises from the toxic insult may, for some chemicals, cause transformation of the renal epithelial cells to produce tumorigenesis. Additionally, it has been suggested that chronic, repeated exposures to some nephrotoxicants cause repeated cycles of cytotoxicity or altered cellular regulation, followed by repair and proliferation. This in turn leads to tumorigenesis through a so-called epigenetic mechanism.

Examination of Figure 7.4 raises the question of how nephrotoxic chemicals produce their toxic insult. There are several possible mechanisms, which are illustrated in Figure 7.5. In general, most nephrotoxic (or just toxic for that matter) chemicals must undergo some form of metabolism or biotransformation to produce metabolites, some of which are called "reactive intermediates," that are the proximate cause of adverse effects in the cell. Reactive intermediates are so named because they are chemically unstable and tend to

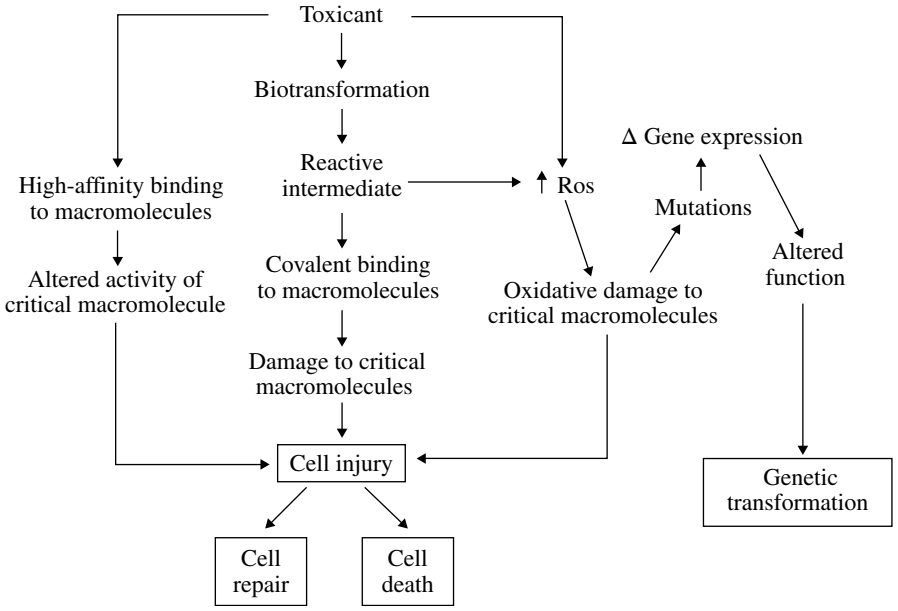


FIGURE 7.5 Scheme of biochemical mechanisms by which nephrotoxic chemicals affect renal cells. For the most part, chemically induced nephrotoxicity is dependent on metabolism or biotransformation to generate a reactive intermediate. Such reactive intermediates may form covalent adducts with cellular DNA, lipid, or protein, thereby altering function, or they may promote or generate reactive oxygen species (ROS). ROS may in turn cause oxidative damage to critical macromolecules, including DNA, leading to altered gene expression and genetic transformation, or may lead to cytotoxicity. As shown in Figure 7.4, the cell injury may lead to cell death or promotion of repair, depending on the extent of the injury and the energy and redox status of the cell.

chemically modify or interact with critical molecules in the cell. These critical molecules include DNA, protein, and lipids. Nephrotoxicants or reactive intermediates derived from their metabolism may also generate partially reduced forms of molecular oxygen (so-called reactive oxygen species (ROS)). ROS are, as the name implies, similar to reactive intermediates in being chemically unstable with a consequent tendency to modify or interact with critical cellular molecules. Similar to the discussion in the previous paragraphs, these interactions may lead to a range of cellular responses, including cell repair, cell death, or mutations that can ultimately result in genetic transformation of the renal cell.

7.3 ASSESSMENT OF RENAL FUNCTION AND INJURY IN THE CLINIC AND IN *IN VIVO* ANIMAL MODELS

One's ultimate goal in treating patients who have been exposed to nephrotoxic chemicals is to monitor kidney function by measurement of various parameters in either urine or blood. These so-called biomarkers are also used in many *in vivo* animal studies to monitor kidney function in an intact organism. Biomarkers can be used for different purposes. For example, parameters that change in response to alterations in kidney function are called *biomarkers of effect*. These are important in quantifying the consequences of toxicant exposures. Besides treatment, we are also interested in prevention of kidney injury. Hence, changes in parameters that occur very quickly after an exposure has occurred but prior to the onset of renal functional changes are termed *biomarkers of exposure*.

Measurement of GFR

One of the primary means of assessing renal function in the whole animal or in patients is to monitor GFR. GFR is defined as the volume of blood completely removed of a substance per unit time (ml/min); GFR = 125 ml/min in healthy, young humans. Two commonly used indicators of GFR are blood urea nitrogen (BUN) or serum creatinine (SCr) levels. Nitrogenous wastes derive from protein and amino acid degradation. Under normal conditions, amino acids are reabsorbed by the proximal tubules and breakdown products of proteins are efficiently excreted, with much of the N-containing waste products being excreted as ammonia or ammonium salts. Under conditions of impaired renal function, however, the efficiency of nitrogenous waste excretion is diminished, leading to increased retention of N-containing compounds. Creatinine ($M_r = 113$) is released from skeletal muscle due to the nonenzymatic hydrolysis of creatine. The rate of creatinine production is generally constant and proportional to muscle mass. In a normal, healthy state, glomerular filtration is nearly complete so that very low levels of creatinine remain in the blood. Using creatinine clearance (C-Cr) as an index of renal function, GFR is calculated by the following equation:

$$\mathrm{C-Cr} = \left(\left(\mathrm{creatinine}\right)_{\mathrm{urine}}\left(\mathrm{mg/ml}\right) \times \mathrm{urine}\left(\mathrm{ml/min}\right)\right) / \mathrm{creatinine}_{\mathrm{serum}}\left(\mathrm{mg/ml}\right).$$

As shown in Figure 7.6, SCr levels typically parallel BUN levels; both parameters generally do not increase significantly until GFR values are decreased below 50% of their optimal levels (i.e., <62.5 ml/min). This fact tells us two important things: First, the kidneys possess a significant reserve capacity, such that fluid and electrolyte homeostasis can be adequately maintained until GFR is reduced to less than 50% of optimum. Second, it is clear that neither SCr nor BUN are particularly sensitive markers of renal function as they typically do not exhibit significant increases until GFR has been reduced by approximately 50%. Because of the potential for variability due to diet and renal secretion, C-Cr data must be carefully interpreted in the clinical setting.

Inulin clearance has historically been the gold standard of exogenously administered markers of GFR. However, its scarcity and high cost have greatly diminished its usefulness in the clinical setting. Inulin is a polymer of fructose with $M_r = 5200$, is inert, does not bind to plasma protein, distributes in the

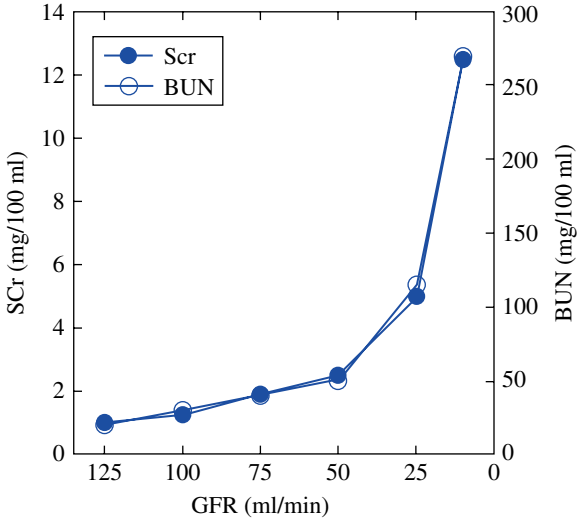


FIGURE 7.6 Variation of serum creatinine (SCr) and blood urea nitrogen (BUN) with glomerular filtration rate (GFR). GFR, which is a standard measure of kidney function and is defined as the movement of water and solutes across the glomerular capillary wall to form an ultrafiltrate of plasma, can be indirectly assessed by measurements of either SCr or BUN. Neither parameter increases significantly until GFR is reduced to less than 50% of maximal capacity. Hence, neither SCr nor BUN is considered a sensitive indicator of renal function. Rather, significant increases in these parameters are indicative of significant renal damage or disease.

extracellular space, is freely filtered at the glomerulus, and is neither reabsorbed nor secreted by renal tubules. It is readily measured in plasma and urine by a colorimetric assay. Glucose is also detected during inulin assay, so it must be determined independently and accounted for in inulin measurements.

GFR can also be estimated by determining the plasma clearance of an intravenous bolus injection of an indicator radionuclide-labeled marker. This is a convenient method that has been used more often than constant infusion or renal clearance techniques. Renal clearance is measured as plasma clearance, or the amount of indicator injected, divided by the integrated area under the plasma concentration curve; models to estimate plasma clearance assume that the volume of distribution and renal excretion are constant over time. Measurements are validated by showing a constant renal excretion, and have been shown for at least two indicators, ^{125}I-labeled iothalamate and ^{51}Cr-labeled ethylenediaminetetracetic acid (EDTA) ($M_r = 292$). Clearance is calculated from the slope and intercept of a line plotted on a logarithmic scale, using the formula Clearance $= V_o(\ln(2))/t_{1/2}$, where V_o is the volume of distribution and $t_{1/2}$ is the half-time for decay in plasma levels; values obtained by this method are multiplied by a constant to correct for systematic errors attributable to overestimation of V_o and a higher concentration of marker in venous than in arterial blood plasma.

Radiocontrast agents were developed to avoid use of radiolabeled compounds. Radiocontrast agents are inherently rich in iodine. Two common radiocontrast agents are *iothalamate* and *diatrizoate meglumine*. Measurements are done by high-performance liquid chromatography and sensitivity allows for as little as 1 ml to be injected s.c. The primary disadvantage of the method is the expense, time, and labor needed to carry out the analyses.

Basic Urinalysis

Urinalysis is a common means to evaluate kidney function in the clinic and in *in vivo* animal models. This typically involves determination of physicochemical properties of the urine (e.g., osmolality or specific gravity, pH, volume), excretion of electrolytes and solutes (e.g., Na^+, K^+, Cl^-, urea), excretion of creatinine as described earlier, excretion of glucose and amino acids, excretion of both high-molecular-weight (>20 kDa) and low-molecular-weight (<20 kDa) proteins that are filtered, and excretion of specific enzymes or other proteins that are released from damaged cells.

Proteinuria is an important indicator of kidney injury. Major high-molecular-weight proteins that are filtered and found in urine of those with damaged kidneys include albumin, IgG, and transferrin; their presence in urine at elevated levels is considered to provide evidence of glomerular injury. In contrast, detection of elevated amounts of low-molecular-weight proteins that are filtered, such as β2-microglobulin, α1-microglobulin, retinol binding protein, and cystatin C, is considered to be indicative of proximal tubular injury.

Secretion of Enzymes and Other Proteins into Urine as Biomarkers of Renal Injury and Site and Mechanism of Action

A large number of enzymes and other proteins can be released from damaged cells or their expression can be increased by chemical exposures and then secreted after injury occurs. Their presence in urine can serve as biomarkers of nephrotoxic effect. Moreover, many of these are released from specific nephron segments, thereby providing information about the site and possibly mechanism of action. A variety of these markers are listed in Table 7.4, which highlights the site of release, species specificity, and comments about sensitivity and utility. Proteins or enzymes detected in urine after a nephrotoxic insult either may arise from damaged cells or may be increased as a result of the toxicant exposure, with the excess being excreted into the urine. Notable among these markers is Kidney Injury Molecule-1 (KIM-1), whose expression may be increased by greater than 50-fold after toxicant exposure and whose appearance in urine precedes that of many other biomarkers.

7.4 *IN VITRO* MODELS TO STUDY RENAL FUNCTION AND INJURY

Although *in vivo* studies are needed to establish the kidneys as a target organ for drugs and other chemicals, mechanistic studies are difficult to conduct in such a complex system. Hence, many investigators have developed numerous *in vitro* models in which to study the mechanism of action of toxicants once the kidneys are identified as a target. Another major use of *in vitro* models is to screen chemicals for their nephrotoxic potency. The primary advantages of this are that *in vitro* models are less expensive than whole-animal studies, they permit the use of paired control and treated samples, they generally provide rapid results, and they do not have the complication of extrarenal effects. An important caveat, however, is that each of these *in vitro* models has limitations that must be taken into account when interpreting results. A general concern with all *in vitro* models is that changes may occur during isolation of the cell or subcellular fraction, thus adding a level of uncertainty to one's ability to use *in vitro* data to make predictions for humans. Additionally, if one is using an *in vitro* model from a laboratory animal species, such as the rat or mouse, an additional level of uncertainty is in extrapolating data from these species to humans. As described elsewhere in this volume, there are often differences between humans and other mammals for which must adjustments must be made in using data for estimating chemical safety and risk.

TABLE 7.4 Urinary Enzymes and Proteins Released or that are Increased in Expression and Secreted from Damaged Renal Cells as Biomarkers of Chemically Induced Nephrotoxicity

Protein or Enzyme	Nephron Segment	Species Specificity	Comments
Fibronectin	Glom, PT (?)	Human, rat	May also arise from plasma and expression may be increased in severe glomerular injury Decreased tubular proteolysis may also contribute
Collagen IV	Glom	Human, mouse, rat	Derived from basement membrane Expression also increased in chronic glomerular injury
GSTA/GSTα	PT	Human, rat	Highest urinary levels with S3 segment–specific toxicants Fairly sensitive index of damage
GSTP	DT	Human	Questionable sensitivity Easily measured by enzymatic or immunoassay
GSTμ	DT	Rat	Questionable sensitivity Easily measured by enzymatic or immunoassay
Clusterin	PT	Rat, monkey	Rapidly increased upon toxicant exposure
Kidney Injury Molecule-1 (KIM-1)	PT	Rat, mouse, human	Biomarker for acute tubular necrosis Used as a urinary biomarker Very sensitive Increased >50-fold upon toxicant exposure
N-Acetyl-β-D-glucosaminidase (NAG)	PT, Glom, papilla	Rat, mouse, dog, monkey, humans	Released from lysosomes of damaged cells Most often used as PT cell marker but not specific Used as a screening marker for chronic exposure to nephrotoxicants in humans
Lysozyme (muraminidase)	PT	Rat, mouse, human	Increased plasma levels also from nonrenal diseases can lead to increased urinary excretion Lysosomal enzyme Highly variable rate of excretion over 24-h period limits usefulness
B-Galactosidase	PT	Rat, mouse, human	Also elevated after glomerular damage Lysosomal enzyme Pronounced diurnal variation in humans
β-Glucuronidase	PT	Rat, mouse, human	Lysosomal enzyme Increased urinary activity in mice after testosterone treatment
Alanine (leucine) aminopeptidase	PT	Rat, dog, human	Brush-border membrane enzyme Appears very early in urine after acute tubular necrosis Some increase in urine due to nonrenal diseases
γ-Glutamyltransferase (GGT)	PT (S3 segment–specific)	Rat, mouse, human	Brush-border membrane enzyme Basal excretion rate higher in male rats
Alkaline phosphatase (intestinal var.)	PT (S3 segment–specific)	Rat, mouse, human	Use as selective S3 damage marker require isoenzyme identification More resistant to shedding from damaged brush-border membrane than other enzymes; indicates more severe damage
Lactate dehydrogenase	DT > PT	Rat, human	Can be released from cytoplasm of any nephron segment but highest in distal nephron
Tamm-Horsfall glycoprotein	TAL	Rat, human	High degree of diurnal variation in rat of excretion Unclear utility as biomarker

DT, distal tubule; Glom, glomerulus; GST, glutathione *S*-transferase; PT, proximal tubule; TAL, thick ascending limb.
?=increased presence in PT unclear.

The advantages, utility, and some specific limitations of each *in vitro* model are summarized here.

Isolated Perfused Kidney

The isolated perfused kidney, most commonly obtained from the rat or rabbit, has the advantages of maintaining intact tissue structure and the ability to distinguish between intrarenal and extrarenal effects, and is useful for studies on a whole-tissue level of processes that require more integrated organ functions. Limitations of this model include its expense, the potential problem of interanimal variability, incomplete definition of conditions, and its short-term viability (typically ≤2 h).

Renal Slices

Renal slices have had considerable utility in drug design and development because of their ease of preparation, the diminished potential for isolation artifacts as compared with other *in vitro* models because of greater retention of intact tubular structure, and their use for drug screening, metabolism, and transport studies. Limitations include relatively short-term viability (≤2 h), the presence of multiple nephron cell types, the potential for poor cellular oxygenation, and limited access to the BBM due to collapsed lumens. Relatively recent advances in slicing technology (e.g., use of so-called precision-cut tissue slicing), which have enabled the preparation of thinner slices, have minimized the problems with poor oxygenation and collapsed lumens.

Isolated Tubules and Tubular Fragments

Isolated tubules are most commonly obtained from rabbit or mouse whereas tubular fragments are more often obtained from rat. Although there is some degree of maintenance of intact tubular structure, digestive enzymes such as collagenase and hyaluronidase are often used to separate out the tubules, thereby introducing the possibility of isolation artifact. Tubules and tubular fragments have several advantages as an experimental model. These include ease of preparation, the ability to obtain highly enriched or purified material derived from specific nephron segments, a high degree of control of incubation conditions, and the ability to perform manipulations with paired controls. They are most often used to study metabolism, acute toxicity, and transport. Limitations include possible damage during isolation and a relatively short lifetime (≤4 h).

Suspensions of Isolated Renal Cells and Primary Cell Culture

Freshly isolated renal cells are another convenient *in vitro* model that possesses all the advantages and limitations of tubules or tubule fragments. Additionally, they permit easy, bidirectional exposure of cells to toxicants but epithelial polarity is lost, which becomes a limitation if the process being studied requires it to be maintained. Single cells in suspension are typically obtained from rat or human kidneys because tubules or tubule fragments are not readily prepared from kidneys of these species. Similar to isolation of tubules or tubule fragments, the most common means of obtaining isolated renal cells is perfusion with digestive enzymes (e.g., collagenase). Although cells can easily be obtained solely from the renal cortex, which are predominantly (>80%) derived from proximal tubules, further enrichment of proximal tubules as well as isolation of other nephron segments can be achieved by the application of methods such as density-gradient centrifugation.

Because of the short lifetime of isolated cells or tubules, many investigators have taken the next step and grown these cells or tubules/tubule fragments in primary culture. By definition, primary culture indicates that the biological material comes directly from the tissue and is grown on a culture surface (e.g., dish or flask) without additional passaging. Passaging is a process by which cells that have grown to confluence (i.e., grown until they form a monolayer that completely fills up the available surface) are removed from their growth surface, diluted with culture medium, and then reapplied to a new culture surface to continue to grow. Because many cellular processes change as a consequence of culture conditions, many investigators prefer to use cells in primary culture as these will exhibit properties and functions closest to those of the *in vivo* tissue. In other cases, however, passaging is necessary to enable longer exposure durations; this is an appropriate or validated biological model if the properties and functions under investigation are maintained during the culture process.

Some basic features of experimental protocols that can be conducted with either freshly isolated renal cells (i.e., acute model) or cell cultures (i.e., chronic model) are summarized in Table 7.5. The types of processes that are studied in each model are determined by the expected responses in each exposure model and the limitations of the model. In the first case, different types of responses will occur with acute, high-dose exposures as compared to those with chronic, low-dose exposures. For the former, parameters such as formation of reactive metabolites, changes in activities of specific enzymes, mitochondrial function, changes in cellular morphology, or various forms of cell death can be quantified. Many other types of responses, however, do not occur on the relatively short time frames in acute exposures; rather, they require longer time periods to develop and be detectable. Some examples include changes in gene expression of regulatory proteins, activities of selected signaling pathways, rates of macromolecular synthesis, or rates of cell proliferation. Many mutagenic and carcinogenic agents that alter cellular proliferation rates may do so, not by causing DNA damage or other direct changes in DNA replication

TABLE 7.5 Experimental Protocols Performed with Primary Cultures for Toxicological Studies

Parameter	Acute Model	Chronic Model
Dose	Low to high	Low
Exposure time	≤24 h	1–10 days
Chemical exposure	Single or multiple	Single, multiple, or continuous
Examples of processes measured	Cellular morphology	Gene expression
	Intermediary metabolism	Signaling pathways
	Enzyme activities	Protein, DNA, RNA synthesis
	Drug metabolism	Mutagenesis
	Oxidative stress markers (e.g., lipid peroxidation, reactive oxygen species, thiol oxidation)	DNA repair
		Cell proliferation
	DNA, protein, or lipid modifications (e.g., adducts, cross-linking)	+
	Mitochondrial respiration, membrane potential	Processes measured in acute model
	Active transport	
	Metabolite concentrations (e.g., ATP, GSH)	
	Lactate dehydrogenase release	
	Trypan blue uptake	
	Apoptosis, autophagy	

ability (genotoxic mechanism), but by changes in signaling pathways or regulatory proteins (epigenetic mechanism). Such mechanisms can be characterized and distinguished in primary cell culture models.

Significant advances have been made in methodology for primary culture of renal epithelial cells over the past 20–30 years. Immortalized cell lines (see next subsection) are usually grown in media containing some form of serum (e.g., fetal bovine serum). This is done to provide the cells with the necessary growth factors and trace elements to which the *in vivo* tissue would be exposed. Most serum preparations, however, are relatively undefined in terms of their chemical composition. Because the cellular material used for primary culture of renal epithelial cells is usually an enriched (as opposed to purified) preparation of the desired cell type, small amounts of contaminating cell types are almost always present. An especially common contaminating cell type is the fibroblast. Two key problems that must be overcome in primary cell culture are: (1) prevention of contaminating cell overgrowth; and (2) maintenance of differentiated cellular function.

If these cell preparations are cultured in the presence of serum, the contaminating cell type such as the fibroblast will typically outgrow the epithelial cell type of interest. To avoid this problem, special serum-free media formulations have been developed that do not contain serum but are supplemented with specific hormones, growth factors, minerals, and other trace elements. These specialized media formulations are termed *serum-free, hormonally defined* media. Typical media supplements for renal epithelial cell culture include insulin, hydrocortisone, epidermal growth factor, transferrin, and selenium. Antibiotics (e.g., penicillin, streptomycin) and an antifungal agent (e.g., amphotericin B) are typically added during the first 2 or 3 days of primary culture to prevent bacterial or fungal growth; these are discontinued after day 3 of culture as they tend to inhibit cell growth.

Use of this serum-free, hormonally defined medium also tends to promote the maintenance of differentiated cellular function whereas the presence of serum usually leads to loss of differentiated function. Examples of differentiated functions or properties of the renal proximal tubule that can be lost during cell culture include the typical cuboidal shape, the high density of microvilli on the luminal or BBM, expression or plasma membrane localization of many transport proteins (e.g., OATs, OCTs, amino acid carriers), expression of Phase I and Phase II drug metabolism enzymes (e.g., CYP, glutathione *S*-transferases), and a high rate of mitochondrial oxidative phosphorylation.

Immortalized Cell Lines

Although primary cell culture has several advantages in terms of reflecting the inherent properties and responses of the *in vivo* cell type of origin, it is a method that also has limitations and constraints to its use. For example, cells must be obtained from the animal or human tissue each time a cell preparation or experimental protocol is initiated. Because of the potential for bacterial or fungal contamination or overgrowth of contaminating cell types, extreme care, including strict adherence to aseptic surgical conditions, must be taken in isolating cells and placing them in culture. There is also the potential for variation from cell preparation to cell preparation. Several immortalized cell lines of renal epithelial origin are commercially available and have been extremely useful in numerous types of studies (Table 7.6). Use of these cell lines has numerous advantages, including ease of use, reproducibility, easy availability, and long lifetime. One use of cell lines that is particularly easier than its performance in

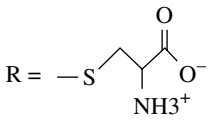
TABLE 7.6 Selected Immortalized, Continuous Renal Cell Lines that are Commercially Available

Cell Line	Species	Presumed Cell Type of Origin
SGE_1	Wistar rat	Glomerulus
NRK-52E	Norway rat	Proximal tubule
$LLC\text{-}PK_1$	Hampshire pig	Proximal tubule
OK	American opossum	Proximal tubule
MCT	Mouse	Proximal tubule
JTC-12	Cynomolgus monkey	Proximal tubule
HK-2	Human	Proximal tubule
GRB-MAL	Rabbit	Medullary thick ascending limb
M-m-TAL-Ic	Mouse	Medullary thick ascending limb
MDCK	Dog	Distal tubule and collecting duct
A6	African clawed toad	Distal tubule and collecting duct

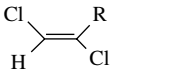
primary cultures is transfection with either cDNAs or small, interfering RNAs (siRNAs). Development of lentiviral or adenoviral vectors for use with primary cell cultures has narrowed this ease-of-use gap between cell lines and primary culture. Finally, as discussed earlier, use of these various cell lines must be in some manner validated for the process under investigation.

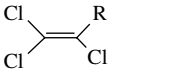
7.5 EXAMPLES OF ENVIRONMENTAL, INDUSTRIAL, AND THERAPEUTIC CHEMICALS THAT PRODUCE NEPHROTOXICITY

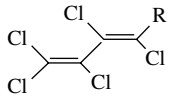
This final section will briefly present selected, well-studied examples of distinct classes of toxicants and therapeutic agents that are selective and potent nephrotoxicants. The examples are meant to illustrate the following principles: (i) some chemicals are selective nephrotoxicants by virtue of the existence of specific transporters that efficiently deliver the chemicals to their target cell; (ii) some chemicals elicit nephrotoxicity after uptake and accumulation in renal cells because of the presence of bioactivation enzymes in the cell; (iii) the mechanism of nephrotoxicity of many chemicals involves effects on specific subcellular organelles (e.g., mitochondria, lysosomes); and (iv) many chemicals can elicit a range of diverse effects in renal cells. These examples are by no means representative of all the different classes of nephrotoxic chemicals. However, they do represent some that are both well characterized and illustrate interesting principles about how specific chemicals act in the kidney to produce toxicity. The reader is referred to selected texts and reviews that are listed at the end of this chapter for other examples of nephrotoxicants and more mechanistic details.



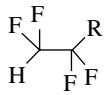
Haloalkenyl *S*-conjugates:

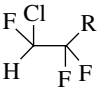

S-(1,2-dichlorovinyl)-L-cysteine (DCVC)


S-(1,2,2-trichlorovinyl)-L-cysteine (TCVC)


S-(1,2,3,4,4-pentachlorobutadienyl)-L-cysteine (PCBC)

Haloalkyl *S*-conjugates:


S-(1,1,2,2-tetrafluoroethyl)-L-cysteine (TFEC)


S-(2-chloro-1,1,2-trifluoroethyl)-L-cysteine (CTFC)

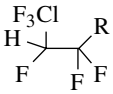

S-(1,1,2,3,3,3-pentafluoropropenyl)-L-cysteine

FIGURE 7.7 Chemical structures of nephrotoxic cysteine conjugates of halogenated solvents. Several halogenated solvents are metabolized by conjugation with glutathione and subsequently to the corresponding cysteine conjugates, which are the penultimate nephrotoxic metabolites. These nephrotoxic cysteine conjugates fall into two classes: haloalkenyl and haloalkyl *S*-conjugates. In both cases, a good leaving group is required for the cysteine conjugate to serve as a substrate for the cysteine conjugate beta-lyase, which generates the ultimate reactive and toxic metabolite.

Halogenated Solvents: Tri- and Perchloroethylene

Halogenated solvents represent a broad class of chemicals that are both widely used in numerous industrial applications and are major environmental contaminants. Those halogenated solvents that produce selective and potent nephrotoxicity have some features in common (Figure 7.7): (i) the nephrotoxicity is dependent on conjugation with glutathione (GSH); (ii) further metabolism to the cysteine conjugate generates the penultimate, nephrotoxic species; and (iii) the portion of the cysteine conjugate derived from the parent molecule provides a good leaving group for the subsequent reaction that generates the ultimate, nephrotoxic metabolite.

GSH is an important component of cellular antioxidant defense and is typically considered protective against most cytotoxic chemicals (e.g., see discussion later about APAP). In the case of halogenated solvents such as trichloroethylene, perchloroethylene, hexachlorobutadiene,

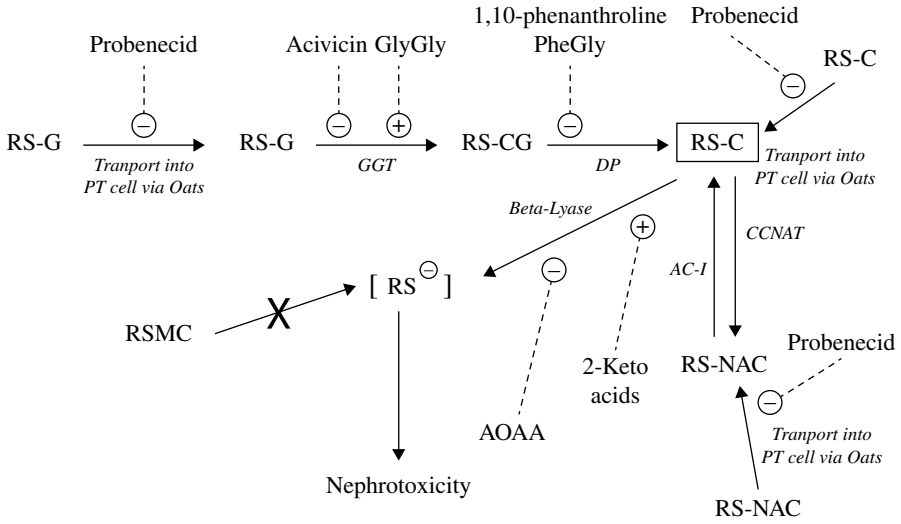


FIGURE 7.8 Experimental approaches to validation of the bioactivation pathway for nephrotoxic halogenated solvents. The reaction scheme illustrates experimental approaches that have been used in both *in vivo* animal studies and various *in vitro* cellular and subcellular models to validate the steps by which glutathione (GSH) conjugates of various halogenated solvents are bioactivated in renal proximal tubular cells. –, inhibition; +, stimulation. AC-I, acylase-I; Acivicin, L-(αS,5S)-α-amino-3-chloro-4,5-dihydro-5-isoxazoleacetic acid; AOAA, amino-oxyacetic acid; CCNAT, cysteine conjugate N-acetyltransferase; DP, dipeptidase; GGT, γ-glutamyltransferase; Oat, organic anion transporter; RS-C, cysteine conjugate; RS-CG, cysteinylglycine conjugate; RS-G, glutathione conjugate; RSMC, *S*-methylcysteine conjugate; RS-NAC, N-acetylcysteine conjugate (mercapturate).

and tetrafluoroethylene, however, it promotes and is required for the bioactivation of these chemicals. For these chemicals, therefore, GSH is critical for producing toxicity and is unique and distinct from its normal function in protecting cells from various drugs and chemicals. A schematic of how each step in the GSH-dependent nephrotoxicity of these solvents has been validated is shown in Figure 7.8. It should be emphasized that these validations have been accomplished both with *in vitro* renal models and with multiple *in vivo* animal models. This schematic clearly illustrates the importance of membrane transport and metabolism in nephrotoxicity, by showing that inhibition of OA^- transporters with probenecid diminishes toxicity whereas inhibitors or cosubstrates of each metabolic reaction either diminish or exacerbate, respectively, toxicity. Additionally, substitution of a methylated cysteine conjugate substrate (which cannot undergo the ultimate bioactivation reaction due to the methyl group) prevents toxicity.

Heavy Metals: Mercury and Cadmium

Heavy metals such as mercury (Hg) and cadmium (Cd) are important environmental pollutants that produce significant renal injury among other target organ effects. Although some differences exist in potency, pharmacokinetics, and molecular forms to which the kidneys may be exposed to either of these metals, some important common features do exist that underlie the nephrotoxicity of these and other heavy metals as well, such as arsenic and lead. For example, both Hg and Cd can bind to and induce the synthesis of metallothioneins (MT), which are a family of small ($M_r = 12$ kDa), cysteine-rich proteins that are important in metal ion metabolism and disposition. Both metal ions primarily bind to sulfhydryl groups such as those on cysteine, GSH, or cysteinyl residues of proteins such as albumin. Another similarity is that both Hg and Cd can be transported into renal proximal tubular cells as complexes with low-molecular-weight thiols. These thiol–metal complexes function as molecular mimics of endogenous thiol- or disulfide-containing compounds and are substrates for various amino acid and OA^- transporters. In terms of their mechanism of action, both similarities and differences exist. For example, both Hg and Cd can increase intracellular production of ROS and target mitochondria within the proximal tubular epithelial cell. One mechanistic difference is that proximal tubular cells are often exposed to Cd as a complex with MT; the Cd–MT complex can be accumulated within the cell by endocytosis, after which it specifically accumulates within lysosomes.

With regard to speciation of the metal ions, Cd is typically presented to the proximal tubular cell in its ionic form as either a complex with a protein (e.g., MT or albumin) or a low-molecular-weight thiol (e.g., GSH, L-cysteine, or *N*-acetyl-L-cysteine). While Cd complexes do accumulate in the kidneys and liver, at least acutely, chronic exposure to Cd results in its accumulation in bone as the metal displaces and substitutes for calcium ions. In contrast, Hg can exist in either inorganic forms (i.e., mercuric or mercurous salts) or in

organic forms (e.g., methylmercury). Whereas inorganic forms of Hg accumulate predominantly in the kidneys, organic forms accumulate in both the kidneys and the central nervous system, the latter being responsible for its neurotoxicity.

Analgesics: Acetaminophen

APAP is an analgesic (commercial name Tylenol®; also called paracetamol) that is considered safe in normal doses. Unfortunately, APAP is subject to abuse and is a prominent cause of poisonings, resulting primarily in liver injury and secondarily in kidney injury. As with most other chemicals, APAP toxicity is dependent on its metabolism to reactive species (Figure 7.9). The initial step of APAP metabolism is catalyzed by either CYP or by either of the Phase II conjugating enzymes uridine diphosphate (UDP)-glucuronosyltransferase or sulfotransferase. Metabolism to the glucuronide or sulfate conjugate results in detoxification, with these highly polar metabolites being readily excreted in urine. In contrast, oxidation by CYP results in formation of a highly reactive quinoneimine (shown in brackets) that can result in cytotoxicity if it reacts with proteins. Under normal conditions and in the absence of an overdose of APAP, the quinoneimine reacts with GSH, eventually forming a mercapturate that is excreted in urine. In cases of APAP overdose or when GSH content is depleted, the reactive metabolite can bind to proteins causing toxicity. In acute overdose, these reactions occur in the renal cortex (primarily

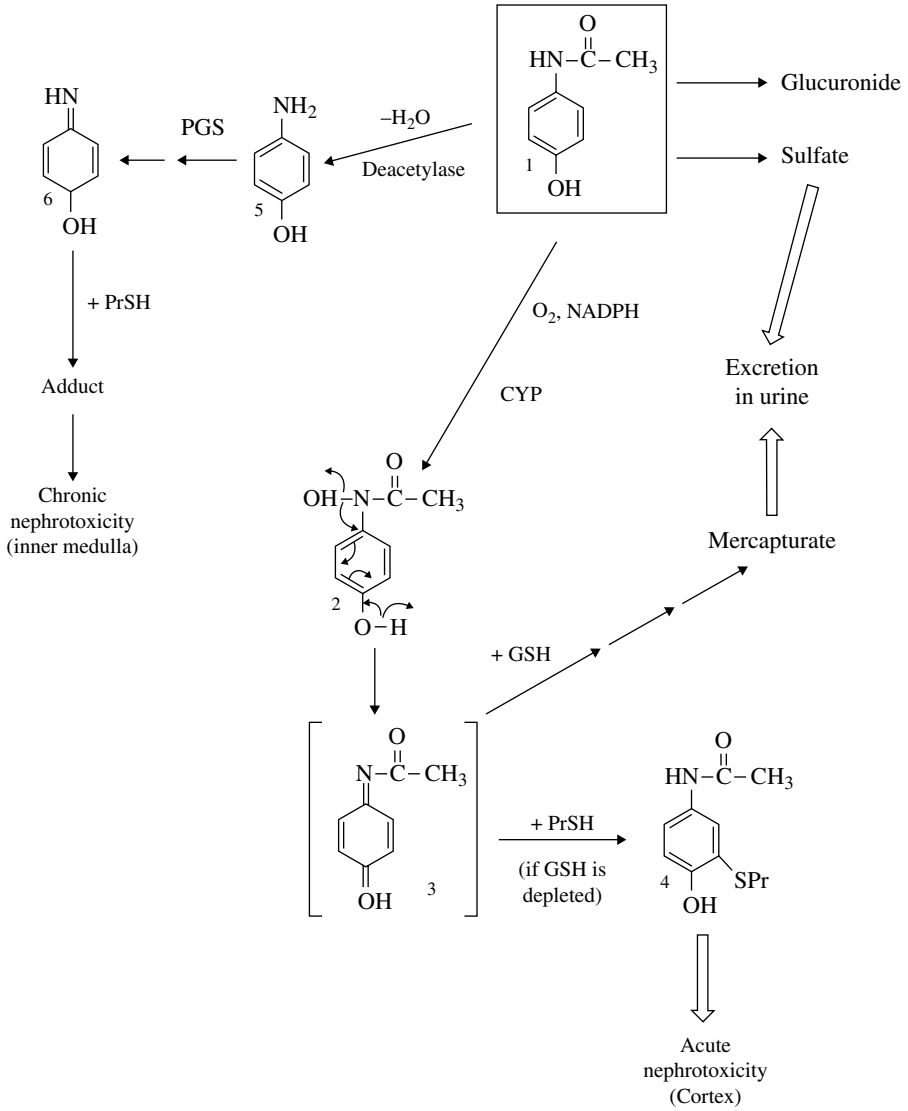


FIGURE 7.9 Metabolic pathways for detoxification and bioactivation of acetaminophen (APAP). Metabolites: 1, APAP (parent compound); 2, N-OH APAP; 3, *N*-acetyl-*p*-benzoquinoneimine (NAPQI; shown in brackets because it is a chemically unstable, reactive intermediate); 4, APAP adduct with a protein; 5, *p*-aminophenol (PAP); 6, *p*-benzoquinoneimine. CYP, cytochrome P450; GSH, glutathione; PGS, prostaglandin synthetase; PrSH, protein sulfhydryl group.

the proximal tubules), leading to proximal tubular necrosis. In chronic APAP abuse, however, the patterns of renal blood flow lead to accumulation of the deacetylated form of APAP (*p*-aminophenol; metabolite 5) in the renal inner medulla, resulting in oxidation by prostaglandin synthetase (PGS) and a form of nephrotoxicity called papillary necrosis.

Antibiotics: Cephalosporins

Cephalosporin antibiotics, which are structurally and mechanistically related to penicillins (Figure 7.10), are commonly used in treatment of Gram-negative bacterial infections, although the original forms were active against Gram-positive bacterial infections. While most of the cephalosporins that are in current therapeutic use are safe and effective, the utility of some of the earlier congeners (so-called first-generation cephalosporin antibiotics) was limited by nephrotoxicity. Some of these first-generation cephalosporins, such as cephaloridine (CPH), are actively accumulated by renal proximal tubular cells (cf. Figure 7.2, pathway C), leading to high intracellular concentrations and cytotoxicity.

CPH represents an interesting example of this class of agents that illustrates the importance of both transport and metabolism in nephrotoxicity. The mechanism of CPH-induced cytotoxicity in renal proximal tubular cells appears to be multifaceted, involving potential bioactivation reactions on three regions of the molecule (Figure 7.11). These three regions of CPH each undergo a fundamentally different metabolic pathway: the thiophene ring serves as a site for CYP-dependent epoxide formation; the central cephem ring (which contains the beta-lactam ring) undergoes a

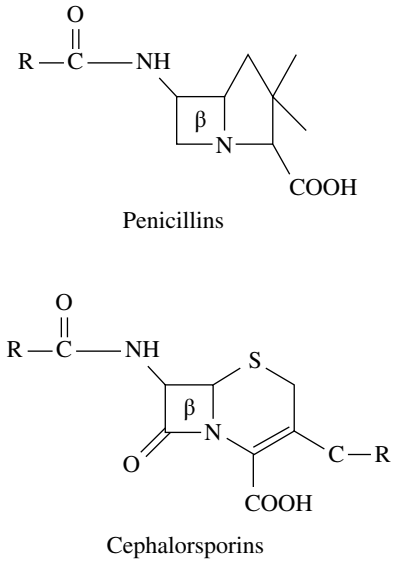


FIGURE 7.10 Comparison between chemical structures of penicillins and cephalosporins. Both classes of antibiotics share a beta-lactam ring attached to a heterocyclic ring, which is a five-membered ring for penicillins and a six-membered ring for cephalosporins.

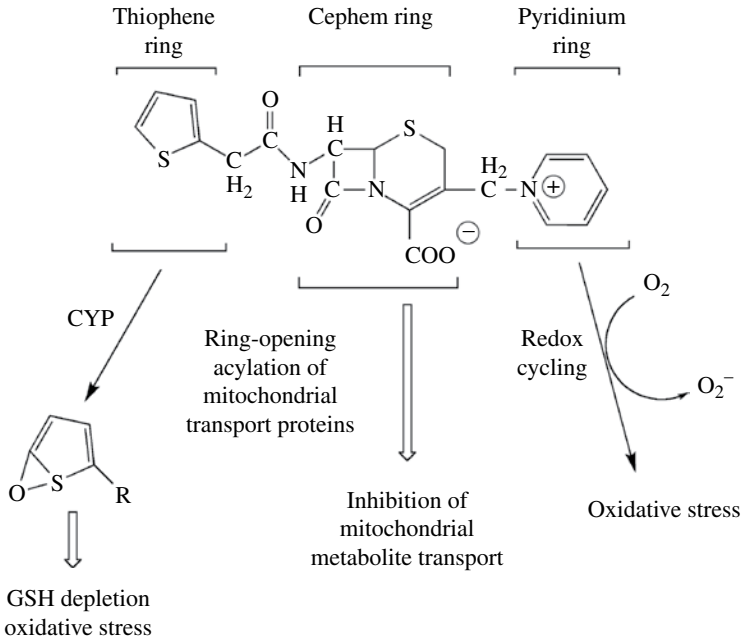


FIGURE 7.11 Proposed bioactivation mechanisms for cephaloridine (CPH). The first-generation cephalosporin antibiotic CPH consists of three structural regions, a thiophene ring, a cephem ring, and a pyridinium ring. The thiophene ring is believed to be a potential substrate for cytochrome P450 (CYP), which generates a reactive epoxide that can cause glutathione (GSH) depletion and oxidative stress. The cephem ring, which is the structural feature common to all cephalosporin antibiotics, has been shown to undergo a ring-opening reaction that results in acylation of proteins, particularly in renal mitochondria. The pyridinium ring, which is analogous to the herbicide paraquat, is thought to be capable of undergoing reversible oxidation and reduction reactions with molecular oxygen, thereby generating superoxide anion.

ring-opening reaction, generating reactive metabolites that acylate mitochondrial transporters, leading to inhibition of their function; the pyridinium ring, which is similar in structure to that of the herbicide and lung toxicant paraquat, can undergo what is termed redox cycling, generating ROS and oxidative cell injury. Redox cycling is a process by which a chemical that can exist in two oxidation states reacts nonenzymatically with molecular oxygen, generating superoxide anion, and reversibly alternates between the two states.

Antineoplastics: Cisplatin

CDDP is a well-known antineoplastic agent whose utility is limited by nephrotoxicity. Its mechanism of action against tumor cells is to bind to and cross-link DNA, thereby inhibiting DNA replication and ultimately leading to cell death by apoptosis. The dose-limiting nephrotoxicity appears to occur by both a genotoxic (i.e., DNA-altering) and oxidative stress mechanism. Once accumulated by the proximal tubular cell, CDDP loses its chloride ions by hydrolysis and either interacts with cellular DNA in the nucleus (similar to its mechanism of action in tumor cells) or inhibits mitochondrial function, thereby generating ROS.

7.6 SUMMARY

The kidneys are frequent target organs for toxic drugs and other chemicals because of their unique physiology and biochemistry. Major determining factors for this susceptibility to chemically induced injury are renal blood flow, the urinary concentrating mechanism, the processes of filtration, reabsorption, and secretion, and the presence within renal epithelial cells (particularly those of the proximal tubule) of bioactivation enzymes. This chapter also reviewed methods to assess renal function, including those that can be used in the clinical or *in vivo* setting, and described the key properties of various *in vitro* models with an emphasis on the key uses and advantages as well as the limitations of each model. Finally, examples of selected nephrotoxicants that represent different classes of chemicals that act by distinct mechanisms were highlighted to further emphasize and illustrate the importance of transport and metabolism in determining renal cellular injury.

SUGGESTED READING

Bandara LR, Kelly MD, Lock EA, Kennedy S. A correlation between a proteomic evaluation and conventional measurements in the assessment of renal proximal tubular toxicity. Toxicol Sci 2003;73:195–206.

Klaassen CD, editor. *Casarett & Doull's Toxicology*. 6th ed. New York: McGraw-Hill; 2001.

Lash LH. In vitro methods of assessing renal damage. Toxicol Pathol 1998;26:33–42.

Lash LH. Renal toxicology: cytoprotective systems within the kidney. In: Schnellmann RG, editor. *Comprehensive Series in Toxicology, Kidney Toxicology*. Volume 7. 2nd ed. Oxford: Elsevier; 2010. p 117–150.

Lash LH, Cummings BS. Mechanisms of toxicant-induced acute kidney injury. In: Schnellmann RG, editor. *Comprehensive Series in Toxicology, Kidney Toxicology*. Volume 7. 2nd ed. Oxford: Elsevier; 2010. p 81–115.

Robertson EE, Rankin GO. Human renal organic anion transporters: characteristics and contributions to drug and drug metabolite excretion. Pharmacol Ther 2006;109:399–412.

Sanz AB, Santamaría B, Ruiz-Ortega M, Egido J, Ortiz A. Mechanisms of renal apoptosis in health and disease. J Am Soc Nephrol 2008;19:1634–1642.

Schrier RW, Gottschalk CW, editors. *Diseases of the Kidney*. 8th ed. Boston: Little Brown and Co.; 2005.

Sekine T, Miyazaki H, Endou H. Molecular physiology of renal organic anion transporters. Am J Physiol 2006;290:F251–F261.

Tarloff JB, Lash LH, editors. *Toxicology of the Kidney*. 3rd ed. Boca Raton: CRC Press; 2005.

Wright SH, Dantzler WH. Molecular and cellular physiology of renal organic cation and anion transport. Physiol Rev 2004; 84:987–1049.

8

NEUROTOXICITY: TOXIC EFFECTS ON THE NERVOUS SYSTEM

William M. Caudle and Gary W. Miller

Within this chapter we will provide:

- A general overview of the gross anatomical structures of the brain as well as the cellular components that comprise the brain and how they function and contribute to the normal workings of the nervous system.
- Mechanisms of action of specific neurotoxicants or classes of neurotoxicants on various aspects of the nervous system explaining how they can disrupt normal central nervous system (CNS) functions and alter our behaviors.
- A description of how knowledge of neuroanatomy and toxicology are paired to evaluate and diagnose the effects of neurotoxicants, either in the workplace or throughout our daily lives.
- A brief appraisal of various techniques utilized in a laboratory setting that assist the researcher in more effectively evaluating the neurotoxicological potential of various environmental contaminants.

The human nervous system is exquisitely complex, composed of billions of neurons and an even greater number of connections among these neurons, all with the fundamental goal of carrying out many basic functions such as the ability to walk, enjoy the smell of a flower, or carry on a conversation with friends. However, the complexity of these finely orchestrated interactions between neurons within different parts of the brain that are imperative for normal neural functioning also makes the human nervous system extremely vulnerable to the effects of different environmental toxicants, especially during stages of neurodevelopment. Neurodevelopment is composed of several different stages, each critical to the proper generation and connection of cells in the nervous system. These critical periods of growth are exquisitely prescribed cellular and molecular sequence of events that must be allowed to occur for the proper placement of anatomical components and connections to be made that will define the subsequent normal function of the CNS. Adding a layer of complexity to this process is the fact that the developing nervous system is tremendously sensitive to exposure to toxic compounds, which can alter multiple processes of neurodevelopment and have deleterious effects on subsequent neurological function and behavior. Therefore, understanding how the CNS develops is essential not only to appreciate the complex anatomy of the mature brain but also to comprehend how traumatic events or exposure to neurotoxic compounds can impact the development and function of the CNS.

All neural tissue and cellular components of the CNS (neurons and glia) arise from the neural tube in a process called neural induction. Whether these cells become neurons, astrocytes, or oligodendrocytes is determined by their interaction with specific molecular cues that act on receptors located on the neural precursor cells. Interaction with these receptors modulates the expression of a series of genes involved in differentiating these cells into neurons or glia. Disruption of these initial processes can result in significant nervous system malformations, including spina bifida and anencephaly.

Occurring in parallel with the induction of cellular diversity is the formation of the neural tube into the basic anatomical structures of the brain and spinal cord. As seen in neural induction, the activation of particular signaling pathways

Principles of Toxicology: Environmental and Industrial Applications, Third Edition. Edited by Stephen M. Roberts, Robert C. James and Phillip L. Williams.

helps to guide the generation of rudimentary structures that will eventually give rise to recognizable regions of the nervous system. Similarly, altered development of brain regions can have significant effects on neurodevelopment and function.

Neurons generated in the CNS must move from the site of their initial genesis to a distant region, where they continue to differentiate into mature neural circuits. This migration of neurons to their particular regions in the nervous system is a tightly orchestrated series of events involving the exposure of the cells to specific chemical cues that help to attract them to the proper spot, while at the same time repelling them from areas they do not belong. Once they have reached their specific regions, neurons extend their processes and form synaptic connections with multiple other neurons in order to establish the complex circuitry necessary to carry out normal day-to-day neural processes.

With these various intricacies of the brain it can be appreciated that alteration or damage to even a single aspect of this system can have significant repercussions to the normal function of the brain and the behaviors it is charged with overseeing. Thus, in order to fully comprehend and appraise the effects of particular toxins and toxicants on the nervous system, a comprehensive understanding of the nervous system is essential.

8.1 NERVOUS SYSTEM

The nervous system is separated into two anatomically distinct parts: the CNS and the peripheral nervous system (PNS). Although their structure and location may be different, these two aspects of the nervous system are not functionally delineated. Rather, their functions are highly dependent upon the other and cooperate in an integrated fashion to elicit neurobehavioral responses. With this in mind, we will focus this section on describing the general structures and physiology of the major parts of each system and provide a brief appraisal of the interaction between the two when possible.

Central Nervous System

The major components of the CNS are the brain and spinal cord. The spinal cord sends, receives, and then processes sensory information from skin, muscles, and joints to produce movements in response to these inputs. These functions are integrated and facilitated by the brain via the mass of nerve fibers that course throughout the entire length of the spinal cord providing a conduit through which the brain can elicit a coordinated response to the peripheral information it is constantly receiving.

Like the CNS, the brain can be subdivided into several anatomically distinct, yet functionally integrated regions (Figure 8.1). Moving up from the spinal cord, the first part

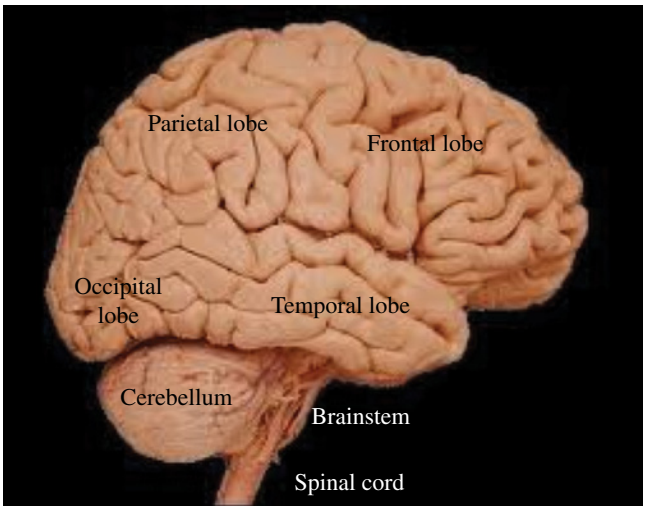


FIGURE 8.1 Superficial structures of the human brain. Picture provided courtesy of the University of Washington School of Medicine, Department of Pathology.

that connects the spinal cord with the brain is the brainstem. The brainstem is involved in general aspects of arousal as well as in acting as a conduit through which nerve fibers between the spinal cord and the rest of the brain travel and communicate. The brainstem can further be divided into specialized regions called the medulla, pons, and midbrain. The medulla and pons participate in the regulation of blood pressure and respiration, while the midbrain plays an important role in controlling the movement of the eyes and body. The brainstem is also important as it is the point from which most of the cranial nerves, which are part of the PNS, emanate. The functions of the cranial nerves will be discussed in greater detail later. Given their role in regulating important aspects of respiration, damage to the medulla and pons could have dire consequences. Indeed, benzodiazepines, such as valium, are CNS depressants, acting on the respiratory centers of the brainstem and suppressing their function, which can result in death.

Continuing further up from the brainstem is the cerebellum. The cerebellum is connected to the brainstem through an intricate series of nerve fibers that arise from the spinal cord and brainstem, as well as other parts of the brain. It receives and integrates information from each of these regions and uses it to regulate vital aspects of posture and balance as we move or stand. Damage to this region can result in a lack of coordination. Thus, it makes sense that many of the uncoordinated movements that are seen following ingestion of ethanol can be attributed to its effect on the cerebellum.

The diencephalon consists of two major regions: the thalamus and the hypothalamus. The thalamus acts as a complex relay station for sensory information that travels from the spinal cord and cranial nerves as well as other parts

of the brain and is routed to the appropriate regions of the brain that will eventually integrate this information and facilitate a motor response. Just below the thalamus is the hypothalamus. The hypothalamus regulates eating, drinking, sleeping, and sexual behavior. In addition, it also works in conjunction with the autonomic nervous system (discussed later) to modulate its function as well as control endocrine processes mediated through the pituitary gland.

The cerebral hemispheres are the most structurally complex and developed regions of the CNS. The cerebral hemispheres consist of four main components: the cerebral cortex, the hippocampal formation, the basal ganglia, and the amygdala. Together, these structures mediate the most sophisticated human behaviors, such as perception, cognition, movement, and emotion. Again, it should be noted that although their anatomical location and structure may be different, each of these regions is intimately integrated, transmitting and receiving information from each other in order to obtain a comprehensive neural image of a stimuli and to fine-tune the response. The most prominent and visually accessible structure of the cerebral hemispheres is the cerebral cortex. The cerebral cortex is most easily identified by its convoluted surface formed by ridges (gyri) and grooves (sulci). This appearance is a result of an evolutionary adaptation that allowed for a greater surface area of the brain to fit within a confined space of the skull. The cerebral cortex can be further divided into four distinct yet integrated lobes. The frontal lobe is essential for motor behavior, cognition, and emotion. In addition, the frontal lobe houses our faculties that allow us to make decisions about what are right or wrong, safe or dangerous. The parietal lobe is concerned with our perception of bodily senses, such as touch and pain, as well as with providing a sense of where our limbs exist in space. In this regard, the parietal lobe works closely with other brain regions involved with motor movement to assist us when we reach for a cup of coffee or hammer a nail. The temporal lobe is involved in the perception of language and sound, in addition to functioning in memory. Finally, the occipital lobe controls the perception and integration of all visual stimuli. As with all other lobes, the occipital lobe is in close reciprocal communication with other regions of the brain in order to refine the sensory information it is receiving as well as the information it transmits to other regions.

Located beneath the lobes of the cerebral cortex are three major components (Figure 8.2). The hippocampus is the part of the brain primarily responsible for learning and memory. Damage to this region can not only affect the ability to recall memories from the past (retrograde amnesia) but also impair the ability to form new memories or learn new tasks (anterograde amnesia). For instance, some algae blooms produce a potent neurotoxicant that targets the hippocampus and destroys vital areas responsible for memory formation and learning. The amygdala mediates our emotions while coordinating our physiological response

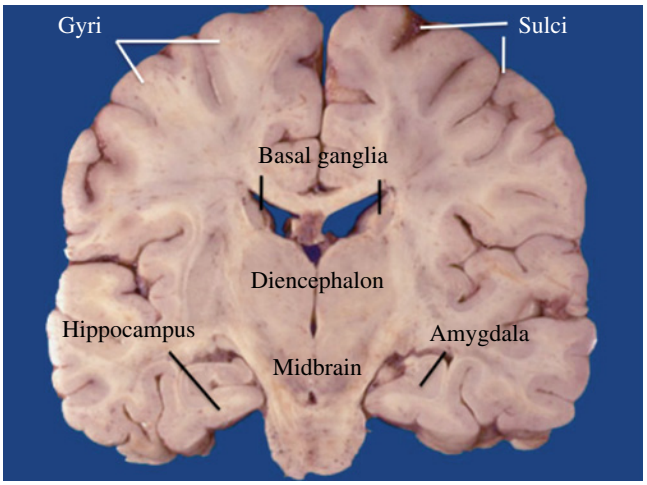


FIGURE 8.2 Coronal aspect of the human brain. Picture provided courtesy of the University of Washington School of Medicine, Department of Pathology.

to stressful or threatening situations. It should be no surprise that the amygdala and the frontal cortex are heavily connected. The amygdala and hippocampus are also tightly integrated in terms of remembering situations or specific events that may have elicited a stressful or threatening response in the past. Damage to the amygdala can result in the inability to appropriately assess a situation as being dangerous or threatening. Finally, the basal ganglia are one of many structures that are involved with modulating movement. Its role is to integrate information with the cerebral cortex and use this to refine or fine-tune our movement. The basal ganglia is the primary structure that is damaged in Parkinson's disease (PD) and is caused by loss of neurons in this region and a reduction in the neurotransmitter dopamine. Interestingly, several neurotoxic compounds, including 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP), carbon monoxide, and carbon disulfide, target the basal ganglia and produce deficits in movement. Further effects of these compounds will be discussed later.

Peripheral Nervous System

The PNS is composed of the nerves that arise from the brain (cranial nerves) and spinal cord (spinal nerves) and are ultimately integrated and modulated by the CNS. Both the cranial and spinal nerves of the PNS can be functionally divided into the somatic and autonomic nervous systems. The somatic division carries motor information to the skeletal muscles and joints as well as transmits sensory information from sensory receptors of the body to the CNS. These functions are usually considered as "voluntary," which is in contrast to the "involuntary" functions of the autonomic division. The somatic nerves are exquisitely sensitive to damage by several classes of neurotoxic compounds,

including gamma-diketones, which can specifically damage these nerves and disrupt the ability to sense touch and pain and conduct motor movements. The autonomic division contains nerve fibers that are responsible for the motor and sensory aspects of internal organs, such as controlling cardiovascular, respiratory, and gastrointestinal functions. The autonomic nervous system can be further divided into the sympathetic and parasympathetic divisions. These divisions maintain an opposing yet complementary function with each other. While the sympathetic division can be considered to be involved in the "fight or flight" physiological response, which serves to enhance blood flow to the heart and muscles, increase respiration, and shunt blood from nonessential biological processes, the parasympathetic division promotes a "rest and digest" response, which lowers blood pressure and breathing, and increases blood flow to organs such as the gastrointestinal tract in order to facilitate these functions.

Neurons

The two major cellular components of the nervous system are neurons and glia. These cells work in conjunction with each other to maintain the overall structure and function of the nervous system. We will consider the structure and function of neurons first. Although several different types of neurons exist, they are all composed of the same four structures, which define their morphology and function as neurons. Dendrites form chemical connections with and receive information from other neurons. This information is passed on via the interaction of chemical neurotransmitters with their specific receptors, which are located on the dendrites. The cell body contains the nucleus, ribosomes, and other basic cellular structures necessary for viability of the neuron. Because other parts of the neurons do not produce many of these necessary proteins and other molecules, they must be transported from the cell body. The most reliable method is via axonal transport, which utilizes specialized cytoskeletal proteins used to provide stabilization of the neuron to move this cargo wherever it is needed. In addition to facilitating axonal transport, the major function of the axon is to transmit information in the form of electrical impulses called action potentials, which are the fundamental signals that carry information from one place to another in the nervous system. These electrical signals are initiated by fluxes in concentrations of potassium and sodium that lie on either side of the neuronal membrane creating an electrical potential, called the resting membrane potential. In general, potassium exists in a greater concentration inside the neuron while sodium has a larger concentration outside the neuron. The action potential is produced through the sequential opening and closing of specific sodium and potassium ion channels that are integrated in the axonal membrane, which allows for a change in ion concentration as well as a change in the electrical potential across the membrane (Figure 8.3). The opening and closing of these channels occurs all the way down the axon as the electrical signal is propagated. There are numerous toxic substances found in nature, such as the poison from a jellyfish sting or poison dart frog secretions that specifically target these ion channels and disrupt their functions.

The speed at which these signals travel is facilitated by the wrapping of myelin around the axon at defined intervals along the entire axon. In the CNS this is done by oligodendrocytes, while Schwann cells are responsible for myelin production in the PNS. Like an electrical wire, the myelin serves as insulation for the axon, which allows the electrical signal to travel down the axon more efficiently. In between each myelinated segment of the axon is a space called the node of Ranvier. Within these intervals where myelin is not present are sodium channels that serve to regenerate the action potential and help in the propagation of the signal down the length of the axon.

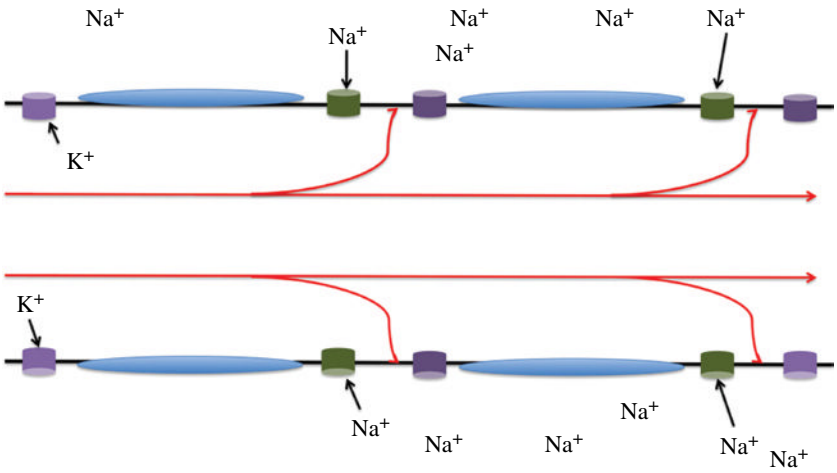


FIGURE 8.3 Diagram of components involved in the neuronal action potential.

When the electrical signal reaches the axon terminal it is transduced into a chemical signal resulting in the release of chemical neurotransmitters. These can be small molecular weight compounds, such as dopamine, norepinephrine, serotonin, glutamate, gamma-Aminobutyric acid (GABA), and acetylcholine as well as larger peptide compounds that also function as neurotransmitters. In general, each of the neurotransmitters is produced within the neuron and is packaged into specialized intracellular compartments called vesicles. When an action potential reaches the axon terminal it signals these transmitter-containing vesicles to fuse with the neuronal membrane and release their contents from the presynaptic terminal. These chemicals then cross a narrow extracellular space called the synapse before they interact with receptors located on the dendrites of the neighboring neuron, also known as the postsynaptic terminal. These receptors have a specific affinity for each neurotransmitter and facilitate the communication of one neuron with another. Depending on the specific neurotransmitter, the binding to the receptors will either result in an activation or inhibition of the target neuron and can generate another action potential, in the case of activation, or can inhibit the initiation of an action potential. Following their interaction with receptors on the postsynaptic neuron the action of the neurotransmitter is terminated through various mechanisms. The most prominent is the reuptake of used neurotransmitters back into the presynaptic neuron through transporters that reside on the neuronal membrane and are specific for each neurotransmitter. An exception to this method is acetylcholine, whose action is terminated within the synapse by the enzyme acetylcholinesterase. Importantly, transporters, receptors, and enzymes involved in the synthesis or degradation of neurotransmitters are primary targets for many well-known toxicants and can disrupt the normal function of the neuron.

Glia

In contrast to neurons, glia do not directly participate in neuronal conduction or information processing. Rather, they serve several metabolic and support functions within the brain. There are two major classes of glia: microglia and macroglia. Microglia serve as the major immune cells of the brain, mediating inflammatory responses following neural damage or infection. The macroglia can be divided into four main types: astrocytes, oligodendrocytes, Schwann cells, and ependymal cells. Astrocytes function in a metabolic and structural support capacity, serving as scaffolding on which newly synthesized neurons traverse on their way to their respective brain regions during neurodevelopment. Oligodendrocytes produce myelin in the CNS, which wraps around the axon and serves to enhance its ability to conduct action potentials. Although Schwann cells serve the same function, they are only found within the PNS where they ensheath spinal nerves that travel to the muscles and visceral organs. The ependymal cells are a specialized cell type that line the fluid-filled cavities of the brain, called the ventricles, and serve a similar function as the epithelial cells of the brain vasculature by forming a selectively permeable barrier and reducing the interaction of compounds found in the blood with the cerebral spinal fluid.

From this general description of the anatomical and cellular components of the brain and spinal cord a sense of the overall complexity and cooperation of the nervous system that is needed for its proper functioning can be appreciated. The neuron is the smallest functional unit of the nervous system and their extensive network of connections with other neurons and other brain regions in addition to the constant receiving, processing, and transmitting of neural information by neurons is what lies at the heart of this complexity. Because of this interconnectedness, disruption in any one part of the neuron can have significant consequences on multiple aspects of nervous function. The neuron itself is exquisitely sensitive to damage through toxic interactions. Indeed, as mentioned earlier, specific parts of the neuron may be targeted by toxic species. However, whether a compound acts upon the cell body or axon or disturbs neurotransmission, the ultimate result is the same: damage to any part of the neuron alters neuronal function and can result in neurobehavioral deficits.

Blood–Brain Barrier

The tremendous metabolic demands of the brain are extensively served by a complex network of vessels that course along the surface of the brain as well as penetrate deep within the brain to serve all regions. This extensive vasculature also leaves the brain highly vulnerable to exposure to toxic compounds that have entered the body and blood stream. As a result, the brain and spinal cord have devised a protective mechanism focused at the level of the vasculature in order to reduce the exposure of the nervous system and its delicate neural components to harmful substances. This restrictive environment is most notably achieved by specialized endothelial cells that compose the walls of the vessels that supply the nervous system.

These cells form tight junctions and contain very few mechanisms for transport of foreign substances from the vessel to the tissue. In addition, vessels are encased by processes that emanate from astrocytes that also help to create a very tight seal around the vessel and maintain a selective control over the movement of constituents from the blood to the neural tissue. Although the major aim of the blood–brain barrier is to keep offending substances out of the nervous system, it is still capable of allowing oxygen, glucose, as well as other pertinent compounds necessary for neuronal viability through the use of specialized transport

mechanisms. These features of the neural vasculature are in distinct contrast to that found in vessels in the periphery, which are fenestrated, relatively leaky, and somewhat promiscuous in what they allow to pass through their walls into the surrounding tissue.

Unfortunately, this system is not flawless and some harmful substances, such as ethanol and cocaine, readily cross the blood–brain barrier, based on their solubility. Indeed, the entry of many toxic compounds is largely dictated by their lipid solubility and their ability to pass through the cells of the vessels. Of note, these compounds have an even easier time gaining access to the fetal and newborn nervous system as the blood–brain barrier is incompletely formed at birth, thus increasing the exposure and vulnerability of the developing brain to toxic insult.

8.2 NEUROTOXICOLOGICAL AGENTS

Neurotoxicological compounds can have differential effects depending upon not only the particular targets and mechanism(s) of action of a toxicant, but also the physiological properties of a specific neuronal population that it acts upon (Table 8.1). Whether a neurotoxicant preferentially affects the cell body, axon, or some aspect of neuronal transmission, the result will be neuronal dysfunction and disruption of a neurophysiological end point subserved by the affected neurons (Figure 8.4). In this section we will describe the mechanism(s) of action of representative compounds on different aspects of neuronal function. Related to this we will appraise the altered neurobehavioral correlates associated with exposure to these compounds and how a greater understanding of the effects of these compounds on the nervous system is evaluated in the laboratory.

TABLE 8.1 Some Common Neurotoxic Compounds

Chemical	Symptom(s)	Sites(s) of Action
Neuronopathy		
Carbon monoxide	Headache, dizziness	Hemoglobin
Carbon tetrachloride	CNS depression, euphoria	Neuronal membrane
Cyanide	Confusion, labored breathing	Mitochondrial cytochrome c
Ethanol	CNS depression, euphoria	Ion channels
Manganese	Parkinsonism, dystonia	Cell body
Phenytoin	Nystagmus, sedation, ataxia	Sodium channels
Axonopathy		
Chloroquine	Peripheral neuropathy, weakness	Lysosomal enzymes
Gold	Peripheral neuropathy, weakness	Axon membrane
Isoniazid	Peripheral neuropathy, sensory	Vitamin B6 depletion
Lithium	Lethargy, ataxia	Nitric oxide signaling
Methyl *n*-butyl ketone	Peripheral neuropathy	Cytoskeletal proteins
Cisplatin	Peripheral neuropathy	DNA bases
Myelinopathy		
Lead	Peripheral neuropathy, tremors	Myelin
Thallium	Seizures, psychosis	Myelin
Tin	Headache, psychosis	Myelin
Neurotransmission		
Amphetamine	Hyperactivity, psychosis	Dopamine transporter
Atropine	Hallucinations, tachycardia	Muscarinic receptors
DDT	Dizziness, convulsions	Ion channels
Deltamethrin	Tremor, seizures	Sodium channels
Nicotine	Excitability, nausea	Nicotinic receptors
TTX	Paresthesia, respiratory failure	Sodium channels

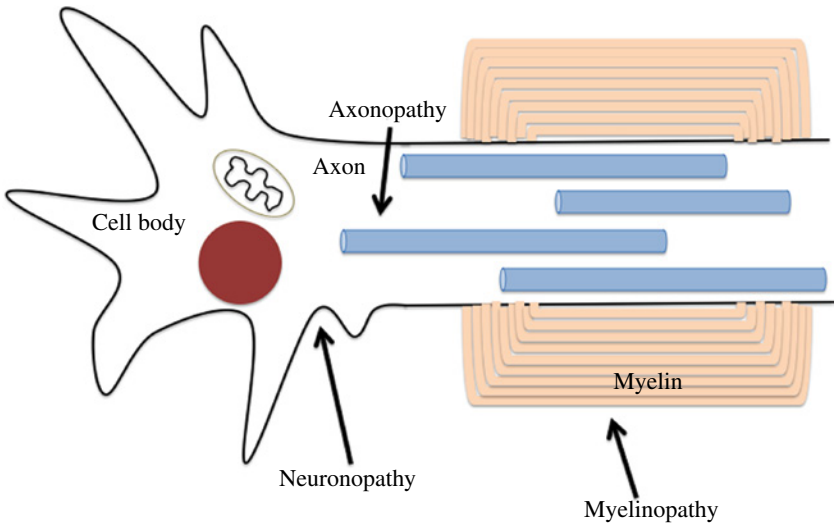


FIGURE 8.4 Sites of neurotoxicological effects.

8.3 NEUROTOXICANTS THAT INDUCE NEURONOPATHY

A shared characteristic of these compounds is that they have a certain predilection for damaging neurons over other cell types. This may be due to certain features of neurons, such as high metabolic rate. Some toxicants are specific for neurons or specific populations of neurons due to particular characteristics of these neurons, which makes them more vulnerable to the toxicant.

Exposure of the human population to methylmercury predominantly occurs through ingestion of contaminated food, mostly through release of industrial waste into bodies of water. Methylmercury is efficiently absorbed by the gastrointestinal tract and can easily make its way across the blood–brain barrier and into the brain via transporters within the barrier. The severity of neurological deficits following methylmercury poisoning is dependent upon the extent of exposure as well as the age of the individual, with young children and the fetus being significantly more susceptible to neuronal loss than adults. Indeed, an in utero exposure to methylmercury can result in severe mental retardation and paralysis, as seen following a widespread contamination of fish in Minamata Bay in Japan. It does not appear that methylmercury poisoning targets a particular cellular component of the neurons. Rather, it tends to disrupt multiple cellular sites resulting in detriments in calcium signaling, mitochondrial function, as well as an accumulation of oxidative stress, all of which can damage the neuron.

As mentioned earlier, the brain is a highly vascularized and metabolically needy structure. Thus, neurotoxicants that affect the efficient delivery of oxygen to brain tissue can have serious consequences on neuronal function. Indeed, exposure to carbon monoxide, which binds to hemoglobin, inhibits its ability to deliver oxygen to tissues. Neurological symptoms of carbon monoxide poisoning include depression, trouble with memory, as well as motor deficits. For reasons that are still unclear, the neuropathological effects of carbon monoxide appear to selectively target neurons of the globus pallidus, which is part of the basal ganglia. However, neuronal losses have also been seen to a lesser extent in the frontal cortex and cerebellum.

MPTP is another compound that attacks and damages the neuronal cell body. The neurotoxic effects of MPTP were identified in the early 1980s following the injection of a synthetic heroin by intravenous drug users that had been accidentally contaminated with MPTP during the drug's synthesis. Within days of injecting the drug, patients developed symptoms of PD with a devastating loss of dopamine neurons in the midbrain, similar to that seen in PD. From a neurotoxicological standpoint, MPTP and its neurotoxic metabolite, MPP+, is quite interesting because instead of having a global neurotoxic effect on multiple neuron populations, it is extremely selective for the dopamine neurons that reside in the midbrain. The specificity of MPP+ for dopamine neurons relates to its ability to be efficiently transported from outside the neuron into the neuron through the dopamine transporter. Facilitating this selectivity is the fact that the dopamine transporter is only present on dopamine neurons and the highest concentration of dopamine neurons is found in the midbrain as well as other parts of the basal ganglia, making these regions and these cells exquisitely more vulnerable to MPTP toxicity.

8.4 NEUROTOXICANTS THAT INDUCE AXONOPATHY

In addition to the cell body, many toxicants target the axon and the cellular components integral to its functioning, creating an axonopathy. Furthermore, these toxicants appear to preferentially damage the long axons of the sensory and motor nerves of the PNS, resulting in sensory and motor deficits that initiate in the arms and legs. Although axonopathies primarily affect the axon while sparing the cell body, damage to the axon creates a holistic dysfunction of the neuron.

n-Hexane exposure, through either occupational exposure or solvent abuse, results in axonal degeneration of the sensory and motor nerves beginning in the hands and feet. With repeated exposures these deficits become more extensive and progress throughout the entire arm and leg. *n*-Hexane is itself not toxic. Rather, its metabolite 2,5-hexadione has been shown to be the intoxicating compound. Evidence suggests that this compound targets the cytoskeletal neurofilaments that are housed throughout the entire length of the axon and function in axonal transport of proteins and other essential cellular components from the cell body to the axon terminal. Impairment of axonal transport of these components significantly compromises the viability of the axon terminal.

Similar to *n*-hexane, carbon disulfide also targets and disrupts the neuronal cytoskeleton, resulting in a peripheral axonopathy that arises in the hands and feet before progressing throughout the limbs with prolonged exposure. Segments of the population most highly exposed to carbon disulfide are those manufacturing viscose rayon, cellophane, and the vulcanization of rubber. Interestingly, carbon disulfide is also a neurotoxic metabolite of the drug disulfiram, better known as Antabuse, which is routinely used to treat chronic alcoholism. Under normal conditions disulfiram functions to inhibit the metabolism of alcohol by acetaldehyde dehydrogenase, which results in a build-up of acetaldehyde in the blood, producing nausea and vomiting and a negative association with alcohol intake. However, excessive exposure to disulfiram has been shown to elicit a dose- and duration-dependent axonopathy similar to that seen with other related compounds.

In addition to industrial sources of axonopathies, several drugs used in a clinical setting to attenuate the growth of various cancerous cells have also been shown to elicit peripheral nerve deficits through their action on the neuronal cytoskeleton. Indeed, colchicine, taxol, and vincristine all exert their anticancer effects through alteration of the cellular cytoskeleton. However, elevated concentrations of these compounds can also target the cytoskeletal proteins of peripheral axons and trigger an axonopathy.

8.5 NEUROTOXICANTS THAT INDUCE MYELINOPATHY

Myelinopathies result from the selective destruction of the myelin that ensheaths the axon and inhibit the conduction of the action potential down the axon. Similar to axonopathies, damage to myelin primarily affects the sensory and motor nerves of the PNS and CNS. Hexachlorophene is a topical disinfectant commonly used in several prescription and over-the-counter formulations. Exposure to hexachlorophene creates edema between the layers of myelin causing them to "split" and become detached from the axon. Diphtheria is a highly contagious bacterial infection of the upper respiratory tract that results in sore throat and fever. Because of its ease of transmission through contact with air-borne secretions from infected individuals, widespread vaccinations have been implemented in most modernized countries, which has effectively eradicated the disease. In addition to causing damage to the heart and liver, the diphtheria toxin can also disturb the axonal myelination of sensory neurons and cause myelinopathy through inhibition of synthesis of the myelin protolipid.

8.6 NEUROTOXICANTS THAT INDUCE ALTERATIONS OF NEUROTRANSMISSION AT THE SYNAPSE

Neurotransmission at the axon terminal can be divided into discrete events such as release of neurotransmitter from the terminal, interaction of the neurotransmitter with the receptors on the postsynaptic neurons, and termination of the neurotransmitter effects through either removal of the transmitter from the synapse or degradation of the transmitter within the synapse. Disruption of any of these events can significantly alter the normal functioning of the neuron and alter the behaviors that are modulated by those neurons (Figure 8.5). Of particular note, most times this alteration in function is transient and occurs without causing explicit damage to the neuron. However, in some instances repeated exposures could result in more permanent neurological

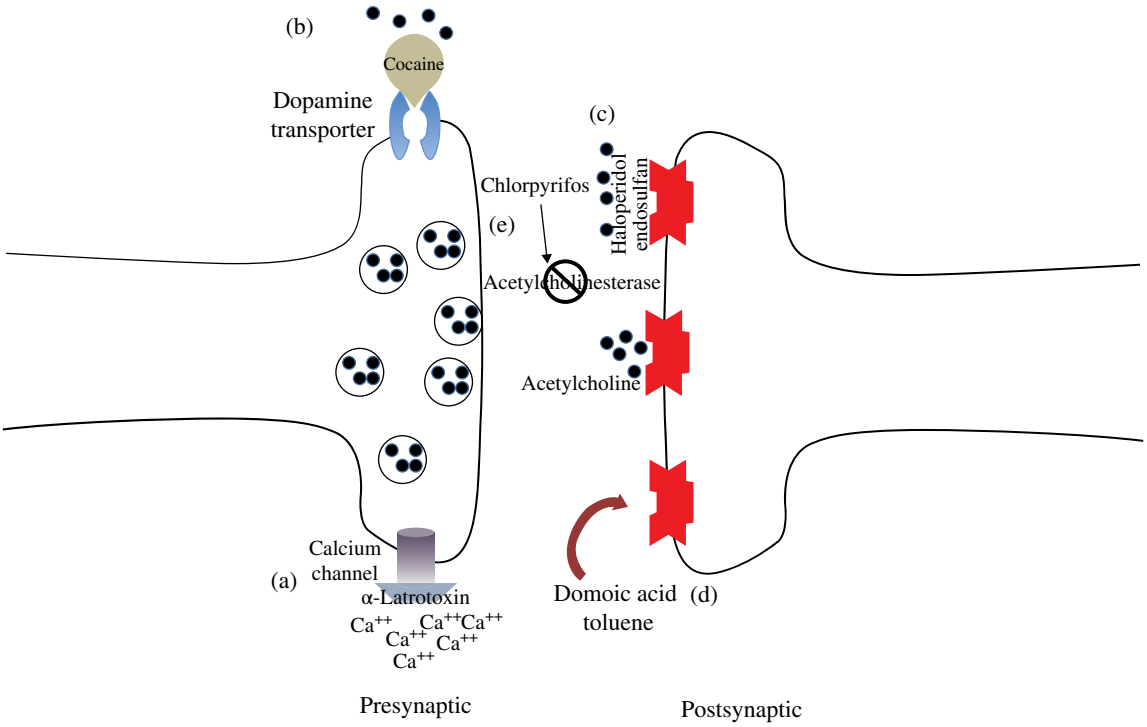


FIGURE 8.5 Sites of altered neurotransmission following toxic insult. (a) Inhibition of Ca^{++} channels by α-latrotoxin. (b) Blockade of the dopamine transporter by cocaine. (c) Antagonism of receptors by haloperidol and endosulfan. (d) Agonism of receptors by domoic acid and toluene. (e) Inhibition of acetylcholinesterase by chlorpyrifos. Note: this is a composite scheme of multiple neurotransmitter systems.

damage and dysfunction. Some of these compounds block specific neurotransmission processes mentioned earlier or even exacerbate or mimic other processes.

α-Latrotoxin is the major neurotoxin found in the widow spider species (black widow). α-Latrotoxin exerts its neurotoxicological effects by targeting the calcium channels located at the presynaptic terminal. In an elegant sequence of events, when an action potential reaches the axon terminal, it triggers the opening of voltage-gated calcium channels, causing an influx of calcium into the terminal. This influx stimulates the neurotransmitter-containing vesicles to bind to the plasma membrane and release their contents into the synapse. Without this calcium signal, the neurotransmitter will not be released. In contrast, a constant supply of calcium would produce a continual release of neurotransmitter. In the case of α-latrotoxin, it facilitates the inward flow of calcium, allowing a flood of calcium and a heightened release of acetylcholine at the muscle, resulting in severe muscle contractions, as well as vomiting and cardiac problems.

A good example of a compound that inhibits an integral aspect of neurotransmission is the blockade of the dopamine transporter by cocaine. The dopamine transporter functions to terminate the action of dopamine by removing it from the synapse and sequestering it back into the presynaptic terminal. By blocking the dopamine transporter, dopamine remains in the synapse where it can repeatedly/continually bind to dopamine receptors and continue to elicit a physiological response. It is believed that this excess exposure to dopamine is what accounts for the euphoric feelings described by cocaine users. While a single use of cocaine may not cause damage to the neurons, repeated or chronic use has been demonstrated to have deleterious consequences.

In addition to the blockade of specific neurotransmitter transporters, some compounds can bind to receptors located on the postsynaptic membrane and can inhibit the normal binding of the neurotransmitter, thus reducing its ability to induce a physiological response. For example, the antipsychotic medication haloperidol is an antagonist for the dopamine receptors in the brain. Blockade of these receptors prevents dopamine from binding and facilitating movement. As a result, one of the major side effects of this treatment are deficits in movement. In addition to pharmaceutical agents, several compounds found in the environment, such as the chlorinated cyclodiene insecticides chlordane and endosulfan, have been shown to block the binding of GABA with its postsynaptic receptors, causing seizures, nausea, and dizziness.

While some chemicals can inhibit the effects of certain neurotransmitters, others can stimulate specific postsynaptic receptors and exacerbate the normal physiological response. For instance, domoic acid is a neurotoxic compound produced and released by algae. Shellfish can take up and accumulate high levels of domoic acid, which is passed on to the human population when we consume shellfish that have been exposed. Domoic acid mimics the effects of the neurotransmitter glutamate, which can then overactivate glutamate receptors and result in severe neurological deficits within the hippocampus and can even cause death. Similarly, toluene, an aromatic hydrocarbon, which is widely abused as an inhalant, binds to the GABA receptors and potentiates the release of GABA in the nervous system. Mild exposure to toluene can cause dizziness, while prolonged exposure can result in unconsciousness and respiratory depression, similar to that seen with benzodiazepines and ethanol.

Not all chemicals perturb neurotransmission through their interaction with the transmitter transporters or receptors. Some act in the synapse between the pre- and postsynaptic terminals. Under normal conditions, acetylcholine is degraded in the synapse by an enzyme called acetylcholinesterase, which functions to terminate the action of the transmitter. Similar to cocaine, inhibition of acetylcholinesterase allows acetylcholine to remain in the synapse and continue to interact with postsynaptic receptors. Exposure to the organophosphate insecticide chlorpyrifos effectively inhibits or inactivates acetylcholinesterase and perpetuates the neurological effects of continuous acetylcholine binding to receptors. Excessive exposure to chlorpyrifos can cause symptoms ranging from salivation and lacrimation to diarrhea, vomiting as well as alterations in movement and consciousness. Interestingly, several of the chemicals that are most well-recognized as being used in biological warfare, such as VX, tabun, and sarin nerve gases, are all acetylcholinesterase inhibitors and were originally manufactured for use as insecticides, but were discontinued after their severe lethality to humans was discovered.

8.7 ROLE FOR GLIA IN NEUROTOXICITY

Although glia do not appear to be direct targets of chemical neurotoxicants, they can play a prominent role in facilitating as well as attenuating the cascade of events that eventually lead to neuronal damage. As discussed earlier, microglia serve as the immune cells of the CNS. These cells are activated under various conditions and can function to assist in maintaining the immune environment. However, under other circumstances, including exposure to toxic compounds or brain trauma, microglia are overactivated. This overstimulation results in the excess production and release of neurotoxic substances such as superoxide, nitric oxide, and tumor necrosis factor-α, which can contribute to neuronal damage. Although microglia can be considered as the predominant contributor to inflammation-mediated neurotoxicity, astrocytes also participate in the neuroinflammatory cascade through their overactivation and release of proinflammatory cytokines.

In contrast to chemical toxicants, glia are directly involved in mediating neuroinflammatory responses following viral

or bacteria infections that result in encephalitis. While general symptoms of encephalitis include fever, headache, and confusion, more severe cases can result in seizures, memory problems, neuronal loss, and death. Bacterial encephalitis can result from parasitic infestations such as toxoplasmosis and malaria, while viral encephalitis can arise from several different origins, including infection with HIV, West Nile, and various strains of influenza. Of the two, viral encephalitis is much more common.

8.8 EVALUATION OF NEUROTOXIC INJURY

As can be appreciated, neurotoxic compounds can affect multiple aspects of the nervous system and cause a myriad of neurological alterations depending on the part of the neuron, the neuronal population, or the brain region affected. Unfortunately, the majority of our knowledge about neurotoxic effects has been obtained following exposures of humans to these species, through either individual incidences, such as excessive exposure of a pesticide applicator to the compound they are spraying, or large-scale accidents like the exposure of workers and the surrounding population following the release of toxic compounds from an industrial source. In these instances it is the clinician that is charged with interpreting the neurological symptoms and delineating the potential neurotoxic cause. Fortunately, clinicians have several tools at their disposal that assist them in these tasks. First and foremost, when a clinician is faced with a patient that they suspect may be a result of contact with a neurotoxic substance a careful narrative of the patients occupational and hobby history needs to be obtained. This is important as it may provide some indication as to the potential for interaction with toxic compounds, such as solvents, pesticides, or metals that could be at the root of the symptoms. While this evaluation may initially involve a questionnaire for the patient, actual quantitative measurements of the environmental levels should be taken if this route of exposure is suspected.

In addition, through biomonitoring, levels of the suspected neurotoxic species or its metabolites can also be measured from the patient's biological fluids (urine, blood) in order to get a quantitative measure of the concentrations of a particular toxic compound within the patient. Biomonitoring is used to assess human exposure to chemical compounds. Using well-defined and sensitive analytical techniques involving gas chromatography/mass spectrometry, levels of the parent compound or its metabolites in various body fluids, including blood, urine, saliva, adipose tissue, and breast milk can be obtained. These measurements can then provide an estimate of the concentrations of a particular compound in the body and can be used to extrapolate possible environmental exposure.

A major utility of biomonitoring is that it allows for a more informed risk assessment of a particular compound or class of compounds that the human population may be exposed to as measurements are conducted using biological samples, thus giving a more relevant assessment. This is in contrast to measuring levels of a compound in the environment (i.e., soil, water, air) or from experimental studies using mice or rats and then attempting to extrapolate human exposure data. In contrast, the lack of measureable levels of a compound does not necessarily suggest that you have not been exposed. Rather, it could imply that the exposure was very small and thus concentrations in the body are below the level of detection or that the body has efficiently metabolized and excreted any measureable concentration of the compound. Finally, as biomonitoring does an excellent job of determining a person's chemical body burden it is not able to reveal the source or the route of the exposure. Thus, other toxicological techniques must be employed in order to provide additional information.

Biomonitoring has routinely been used to evaluate the levels of pesticides in the human population, most notably, the concentrations of the organophosphate insecticides, such as chlorpyrifos and malathion. Although the use of these compounds for household tasks has been banned, they are still readily used in an agricultural setting, with the human exposure occurring through the ingestion of contaminated food. These compounds target and inhibit the function of the enzyme acetylcholinesterase, which normally serves to inactivate acetylcholine at the synapse. This inhibition allows an excess of acetylcholine to react with the receptors and to cause severe neurological reactions in both adults and children. Using a common metabolite of many organophosphate compounds, dialkyl phosphate, investigators have been able to determine the relative levels of organophosphates in the human population. As these compounds are routinely used on our food supply and we continue to be exposed to them, it was determined that approximately 90% of the human population demonstrates some sort of body burden of organophosphate pesticides.

Following a history a thorough neurological exam is necessary in order to gain a comprehensive understanding of the extent and specificities of the patient's neurological deficits. This exam serves to evaluate all neurological functions of the patient, including neuropsychological as well as sensory and motor function. The neuropsychological evaluation covers a broad range of cognitive abilities that are susceptible to neurotoxicological insult, such as general intelligence, memory, language, personality, and mood. Additionally, the patient's motor and sensory capacity is assessed, using behavioral indices such as walking or finger tapping to more quantitative measurements of sensory function using electrophysiological methods to measure nerve conductions and electrical potentials whose deficits may indicate an axonopathy or myelinopathy.

While unfortunate, these observations in human patients can then be utilized by the laboratory researcher in an effort

to better understand the cellular and molecular mechanisms of the neurotoxic effect, so that more effective therapeutics or prevention may be developed. There are multiple techniques that can be used to assess the extent and severity of damage to specific aspects of the nervous system. These approaches range from the use of immortalized cell lines to address cellular and molecular mechanisms of neurotoxicity to whole animal models of toxic exposure that serve to inform us on the neuropathological as well as neurobehavioral end points commonly seen in the human population. Within this section we will briefly appraise multiple laboratory approaches that can be undertaken to address the neurological effects of toxic compounds.

Beginning with an initial characterization of the neurotoxicological effects of a compound in an *in vitro* cell system can yield important information concerning the molecular mechanisms that could be involved in neuronal damage and does so in a relatively quick, inexpensive, and efficient manner, compared with more elaborate animal studies. Within this system, mechanisms of cell death, alterations to the functioning of ion channels, changes in axonal transport, as well as disruption of multiple intracellular components can be evaluated with various techniques. These techniques can be elaborated to a dissociated animal model, in which cells are isolated from the brain or spinal cord of animals and grown in a cell culture plate, similar to an immortalized cell line. Furthermore, the brain of an animal can be cut very thin and used for further evaluation of acute toxicity, either to assess damage or loss of particular proteins or to identify alterations in the electrical properties of specific neuronal circuits. While *in vitro* models provide a wealth of information concerning the neurotoxic characteristics of a compound, these systems can be quite artificial and a poor representation of what is happening in an intact animal.

Animal models of neurotoxicity are very broad and can encompass the use of very simple animal species, such as zebrafish, *Caenorhabditis elegans*, and Drosophila to more evolutionary elaborate animals, such as mice, rats, and nonhuman primates. While each species may be utilized to assess specific aspects of neurotoxicity, for instance, zebrafish have proven advantageous in addressing the neurodevelopmental alterations of some compounds, the predominant animal model used in neurotoxicological research is the rodent. The use of mice and rats, rather than higher species such as monkeys, to evaluate neurotoxicity is very beneficial due to their relative cost, efficiency of treatment, and the similarities of many of their biological functions to that of humans, making it easier to interpret and extrapolate the data to the human population.

Within a whole animal, researchers can gain a more comprehensive understanding of the effects of a neurotoxic species because they are able to evaluate, not only all of the end points discussed earlier in an *in vitro* model, but also several other markers of neuropathology, including neuronal loss, alterations to nerve conduction and impulses within the PNS, and assessment of neurotransmitter release and receptor function. These neuropathological markers can be taken in sum and then extrapolated to the subsequent neurobehavioral outcomes that can also be measured in the same animals. Evaluation of locomotion, cognition, and sensory functions can provide a behavioral correlate to the neuropathological indices that may be observed. An example of the pairing of pathological assessment with behavioral end points is the use of MPTP to induce a parkinsonian phenotype in mice. Following exposure, mice develop many of the same pathological signs, including loss of the neurotransmitter dopamine and dopamine neurons in the substantia nigra pars compacta as well as behavioral symptoms, such as reduced locomotion that are routinely observed in patients with PD. Similar to PD, mice administered the dopamine replacement drug, L-DOPA regain motor function. Finally, with a reliable disease model researchers are able to evaluate the efficacy of various novel therapeutic agents.

8.9 SUMMARY

The human nervous system is exquisitely complex. While this complexity contributes to our functioning in daily life as well as our uniqueness as individuals, it also provides an equally complex series of highly vulnerable targets on which various neurotoxicants can act. From components of the cell body that help generate new and imperative proteins to cellular receptors and transporters involved in neurotransmission, all these parts must work together in synchrony in order for the nervous system to function properly. However, all these parts also have the potential to be targeted and damaged by neurotoxicants. Disruption to any one of these parts, no matter how slight, can have serious consequences on the normal functioning of that neuron, the neuronal circuit it is a part of, and the neurobehaviors it subserves. That said, portions of the human nervous system have a notable reserve capacity than can help buffer against the low level of exposure

To this day, thousands of chemicals exist in our environment with many more being introduced yearly, yet we have very little if any knowledge of how they affect the human brain. Thus, it is essential to continually study how these chemicals affect the nervous system to provide a better understanding of the particular target and mechanism of action of specific chemicals or classes of chemicals. This knowledge will not only assist us in creating potential therapeutic interventions for neurotoxic insults but aid in attenuating or abrogating these insults through the regulation and use of certain compounds and their neurotoxic potential to the human population

SUGGESTED READING

Bruckner JV, Anand SS, Warren DA. Toxic effects of solvents and vapors. In: Klaassen CD, editor. *Casarett and Doull's Toxicology: The Basic Science of Poisons*. 7th ed. New York: McGraw-Hill; 2008.

Carpenter MB. *Core Text of Neuroanatomy.* 4th ed. Baltimore: Williams and Wilkins; 1991.

Chang LW, Slikker W Jr, editors. *Neurotoxicology: Approaches and Methods.* London: Academic Press; 1995.

Costa LG. Toxic effects of pesticides. In: Klaassen CD, editor. *Casarett and Doull's Toxicology: The Basic Science of Poisons*. 7th ed. New York: McGraw-Hill; 2008.

Kandel ER, Schwartz JH, Jessell TM, editors. *Principles of Neural Science*. 4th ed. New York: McGraw-Hill; 2000.

Lemke G, editor. *Developmental Neurobiology.* London: Academic Press; 2009.

Levitan IB, Kaczmarek LK, editors. *The Neuron: Cell and Molecular Biology.* 3rd ed. New York: Oxford University Press; 2002.

Liu J, Goyer RA, Waalkes MP. Toxic effects of metals. In: Klaassen CD, editor. *Casarett and Doull's Toxicology: The Basic Science of Poisons*. 7th ed. New York: McGraw-Hill; 2008.

Moser VC, Aschner M, Richardson RJ, Philbert MJ. Toxic responses of the nervous system. In: Klaassen CD, editor. *Casarett and Doull's Toxicology: The Basic Science of Poisons*. 7th ed. New York: McGraw-Hill; 2008.

Purves D, Augustine GJ, Fitzpatrick D, Hall WC, LaMantia A-S, McNamara JO, White LE, editors. *Neuroscience.* 4th ed. Sunderland (MA): Sinaur Associates, Inc.; 2008.

Watkins JB III. Properties and toxicities of animal venoms. In: Klaassen CD, editor. *Casarett and Doull's Toxicology: The Basic Science of Poisons*. 7th ed. New York: McGraw-Hill; 2008.

9

DERMATOTOXICITY: TOXIC EFFECTS ON THE SKIN

Sailesh Konda and Howard I. Maibach

The skin is the first line of defense against external toxicant exposure. Normal skin is an excellent barrier to many substances, but because of its 1.5–2.0 m^2 surface area, it can act as a portal of entry for many diverse chemicals that come into contact with it, causing dermal and/or systemic effects. Understanding the skin composition and factors that influence percutaneous penetration are prerequisites to understanding the mechanisms and manifestations of toxicant injury of the skin. This chapter capsules:

- Composition of the skin
- The skin's ability to defend against toxicants
- Types of skin maladies
- Commonly used tests to predict chemical hazards and risks

9.1 HISTOLOGY

The skin is composed of two layers: the outer epidermis and the underlying dermis. The layers are firmly associated and together form a barrier that ranges in thickness from 0.5 to 4 mm or more in different body parts. The epidermis and dermis are separated by a basement membrane, which has an undulating appearance. The uneven interface gives rise to dermal ridges and provides the basis for the fingerprints used in personal identification since the patterns of ridges are unique for each individual. Hair follicles, sebaceous glands, and eccrine and apocrine (i.e., sweat) glands span the epidermis and are embedded in the dermis. A third subcutaneous layer lies below the dermis and is composed mainly of adipocytes. Even though this layer is not technically part of the skin, it plays an integral role by acting as a heat insulator and shock absorber (Figure 9.1).

The epidermis, composed of several layers of cells, some living and some dead, is composed of keratinocytes. These cells undergo keratinization, a process during which they migrate upward from the lower epidermis (basal layer) and accumulate keratin (80%, once fully mature and nonviable). By the time they reach the outer layer, the stratum corneum, the cells are no longer viable. They flatten and lose their aqueous environment, which is replaced by lipids. The stratum corneum superficial cells are continuously lost and must be replaced by new cells migrating from the lowest layers. The lower layers immediately adjacent to the dermis (stratum germinativum and stratum spinosum) are responsible for the continual supply of new keratinocytes and initiation of keratinization. Migration and differentiation of keratinocytes from the lower viable layers to the upper stratum corneum take approximately 2 weeks, with another 2 weeks elapsing before the keratinocytes are shed from the surface. The lowest two layers of the epidermis also contain melanocytes, which produce the pigment melanin. Melanocytes extrude melanin, which is taken up by the surrounding epidermal cells, giving them their characteristic brown color. Langerhans cells are also found in these layers and play a role in the skin's immune response to foreign agents.

The dermis has a largely supportive function and represents about 90% of the skin in thickness. The predominant cells are fibroblasts, macrophages, and adipocytes. Fibroblasts secrete collagen and elastin, thereby providing the skin with elastic properties. Dermis is well supplied with lymph and blood capillaries. The capillaries terminate in the dermis without extending into the epidermis. A toxicant must penetrate the epidermis and enter the dermis in order to enter the

Principles of Toxicology: Environmental and Industrial Applications, Third Edition. Edited by Stephen M. Roberts, Robert C. James and Phillip L. Williams.

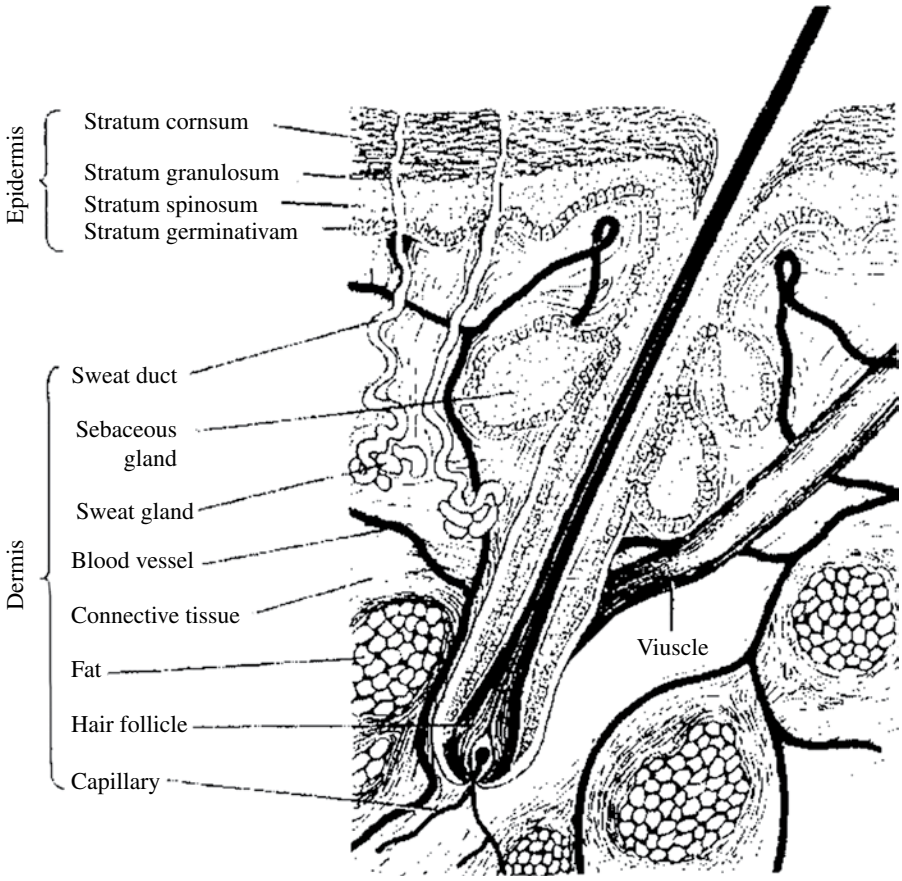


FIGURE 9.1 Diagram of a cross-section of skin. Reprinted with permission Klaassen, eds. (2008). © 2008 McGraw-Hill.

systemic circulation; however, once the stratum corneum is breached, the remaining layers pose little resistance to toxicant penetration. Hair follicles are embedded within the dermis and have a capillary at the bulb of the follicle. In some instances, hair can enhance toxicant absorption across skin by providing a shunt to the blood supply at the base of the follicle. Eccrine glands are embedded deep within the dermis, and coiled sweat ducts wind upward through the epidermis and out through the stratum corneum.

9.2 FUNCTIONS

The skin is an effective barrier to many substances, but is an imperfect barrier to some. This is an important concept because even though relatively small amounts of chemicals cross the skin, they can be sufficient to cause local and/or systemic toxicity. Chemical passage through the skin appears to be by passive diffusion with no evidence so far of active transport. The stratum corneum is the primary layer governing the rate of diffusion, which is slow for most chemicals. This layer also inhibits water loss by diffusion and evaporation from the body except, of course, at the sweat glands, which helps regulate body temperature. The viable layers of the epidermis and the dermis are poor barriers to toxicants, since hydrophilic agents readily diffuse into the intercellular water and hydrophobic agents can embed in cell membranes, eventually reaching the blood supply in the dermis.

Several factors influence the diffusion rate of chemicals across the stratum corneum. In general, hydrophobic agents of low molecular weight can permeate the skin better than those that are hydrophilic and of high molecular weight. This is due to the low water and high lipid content of the stratum corneum, which allows hydrophobic agents to penetrate more readily. However, if the skin becomes hydrated on prolonged exposure to water, its effectiveness as a barrier to hydrophilic substances is reduced. Often, the skin of experimental lab animals is covered with plastic wrap to enhance the hydration of the skin and increase the rate of uptake of agents applied to the skin surface. For compounds with the same hydrophobicity, the smaller compound will diffuse across the skin fastest because its rate of diffusion is quickest. A good example of the diffusion of a class of toxicants across the skin that can cause systemic toxicity is the organophosphate pesticides (e.g., parathion) used in agriculture. These compounds are hydrophobic, are potent, and lead to systemic effects such as peripheral neuropathy (i.e., nerve damage) and lethality after exposure to only the skin.

The property of diffusion of agents across the skin and the reservoir capacity of the skin can be useful in delivering drugs to the systemic circulation over a prolonged period (typically 1–7 days). Transdermal drug delivery using specially designed skin patches is used to deliver nicotine, estradiol, nitroglycerin, and others. This approach provides a steady dose, avoids large peak plasma concentrations from loading doses, and prevents first-pass metabolism by the liver for agents that are sensitive to metabolism such as nitroglycerin.

The rate of diffusion through the epidermis varies among anatomical sites and is not solely a function of skin thickness. In fact, the skin on the sole or palm has a higher rate of diffusion than the skin of the forearm or abdomen, even though it is much thicker. Therefore, skin thickness is not a useful indicator of how much chemical will reach the systemic circulation in a given time. If the skin is wounded, the barrier to chemicals is compromised and a shorter or direct route to the systemic circulation is available because the skin can no longer repel the chemicals. In addition, diseases (e.g., psoriasis) can compromise the ability of the skin to repel chemicals.

Since World War II, extensive methodology has been developed to quantify the rate and extent of percutaneous penetration—in silico, *in vitro*, and *in vivo*—in artificial membranes, animal, and man.

The skin also provides protection against microorganisms and ultraviolet (UV) radiation. Hydrated skin has a greater risk of becoming infected by microorganisms than does dry skin, which is why soldiers in Vietnam often suffered from foot infections. The stratum corneum and epidermis, but primarily melanin pigmentation, provide some protection against UV radiation by absorbing the energy before it reaches more sensitive cells and causes adverse effects such as DNA damage (Table 9.1).

Another important aspect of the skin's barrier function is its ability to metabolize chemicals that cross the stratum corneum and enter the skin's viable layers. Even though the metabolic activity of the skin on a body weight basis is less than that of the liver, it plays a crucial role in determining the ultimate effects of some chemicals. The epidermis and pilosebaceous units contain the highest levels of metabolic activity, which includes Phase I (e.g., cytochrome P450–mediated) and Phase II enzymes (e.g., epoxide hydrolase, UDP glucuronosyltransferase, glutathione transferase).

TABLE 9.1 Defense Roles of the Skin

Prevent water loss
Act as a barrier for physical trauma
Retard chemical penetration
Prevent ultraviolet light penetration and damage
Inhibit microorganism growth and penetration
Regulate body temperature and electrolyte homeostasis

Some chemicals that cross the skin are simply degraded and eliminated as innocuous metabolites. For others such as benzo(a)pyrene or crude coal tar (the latter is often used in dermatological therapy), metabolism of the parent compound can produce a metabolite that is a putative carcinogen. In addition to metabolizing foreign agents, the skin also has anabolic and catabolic metabolic activity important to its maintenance.

9.3 TYPES OF DERMAL TOXICITY

Irritant Dermatitis

Irritant contact dermatitis is a common occupational disease. The highest incidence of chronic irritant dermatitis of the hands occurs in food handlers, janitorial workers, construction workers, mechanics, metal workers, horticulturists, and those exposed to wet working environments, such as hairdressers, nurses, and domestic workers. Irritant dermatitis is confined to the area of irritant exposure, and because it is not immunity-related, it can occur in anyone given a sufficient chemical exposure. Previous exposure to the chemical is not required to elicit a response as is needed for allergic contact dermatitis because contact irritant dermatitis is not a hypersensitivity reaction (discussed in Section "Allergic Contact Dermatitis"). A range of responses can occur after exposure to an irritant, including, but not limited to, skin reddening (erythema), vesicles, blistering, eczemas or rashes that weep and ooze, hyperkeratosis (thickening of the skin), pustules, and dryness and roughness. Unlike with corrosive chemicals (e.g., strong acids and bases), the ultimate skin damage from most irritant contact dermatitis is not due to the primary actions of the chemicals but to the secondary inflammatory response elicited by the chemical. Note that even though the ultimate inflammatory response elicited by different chemicals may appear the same, they often occur through different mechanisms.

An array of factors influences the ability of an irritant to elicit an inflammatory response. As noted earlier, factors affecting skin permeability and chemical composition of the irritant determine the rate of percutaneous penetration and how much chemical reaches the viable layers of the skin. Other factors determine whether irritant dermatitis occurs and to what magnitude. Higher concentrations and greater amounts of a given agent contacting the skin surface are more likely to elicit a response than lower concentrations and smaller quantities. The genetic makeup and age of the individual plays a critical role in the irritant sensitivity to a particular agent—the same chemical can cause no response in one individual and a dramatic response in another. However, specific genetic factors influencing sensitivity are unknown. In general, infants appear to be more and the elderly less susceptible to irritants. Concomitant disease

TABLE 9.2 Potential Inducers of Irritant Contact Dermatitis

Agent	Examples
Water	—
Cleansers	Soaps and detergents
Bases	Epoxy resin hardeners, lime, cement, and ammonium
Acids	Hydrochloric acid and citric acid
Organic solvents	Many petroleum-based products
Oxidants	Peroxides, benzoyl peroxide, and cyclohexanone
Reducing agents	Thioglycolates
Plants	Orange peel, asparagus, and cucumbers

Source: Adapted from Rietschel (1985).

may increase or decrease sensitivity to an irritant. Extremes in temperature, humidity, sweating, and occlusion can alter the threshold of irritation for a given compound.

The range of agents that can cause irritant dermatitis is extensive and diverse, and all cannot be covered in this section. Table 9.2 lists some commonly encountered classes. All have the potential of causing irritation on primary exposure; however, in the workplace, exposure to a potential irritant often occurs repeatedly and to relatively low quantities. Since the response is dependent on the amount of irritant to which the individual is exposed, repeated exposure (cumulative irritation) may be required before clinical signs of dermatitis appear. Management of irritant dermatitis is based on reducing or avoiding the amount of irritant exposure. Wearing gloves to provide protection against wetness or chemicals and minimizing wet working conditions and hand washing can be helpful. Complete healing of lesions may take several weeks, and the likelihood of a flare-up is often increased for months. Controlled studies failed to document efficacy of topical corticosteroids and, in fact, they may exacerbate the condition. Tap water compresses are traditionally utilized and were effective in one controlled study.

Extremely corrosive and reactive chemicals can cause immediate coagulative necrosis at the site of contact resulting in substantial tissue damage. These chemicals differ from those that cause irritant contact dermatitis in that they cause nonselective damage at the site of contact that is not a result of the secondary inflammatory response. They cause damage resulting from their reactivity, such as acids precipitating proteins and solvents dissolving cell membranes, both resulting in cell damage, death, or disruption of the keratin ultrastructure. The resulting damage is in direct proportion to the concentration of chemical in contact with the tissue. Note that such irritants are not always in a liquid form. Many are solid chemicals that become hydrated on contact with the skin, and gaseous agents are often converted to acids on contact with water available on the skin and mucous membranes. Ammonia, hydrogen chloride, hydrogen peroxide, phenol, chlorine, sodium hydroxide, and some antiseptic or germicidal agents (e.g., cresol, iodine, boric acid, hexochlorophene, thimerosal) are some of the commonly encountered primary irritants that can cause skin burns.

Some chemicals can cause ulceration of the skin. This involves sloughing of the epidermis and damage to the exposed dermis. Ulcers are commonly triggered by acids, burns, and trauma, and can occur on mucous membranes and the skin. A commonly encountered compound that induces ulcers is the chromate in occlusion-related cement exposure.

Predictive Testing Common sense and experience dictate that predictive animal and human testing can decrease the risk of irritant dermatitis.

Allergic Contact Dermatitis

Allergic contact dermatitis is a delayed type IV hypersensitivity reaction mediated by a triggered immune response. Typical of a true immune reaction, minute quantities of the allergenic agent can trigger a response. This differentiates it from irritant dermatitis, which typically requires a higher dose/area. Allergic contact dermatitis can be similar to irritant contact dermatitis clinically, but allergic contact dermatitis tends to be more severe and is not always restricted to the body part exposed.

On first exposure to the allergenic chemical, little or no response occurs. After this first exposure, the individual becomes sensitized to the chemical, and subsequent exposures elicit the typical delayed type IV hypersensitivity reaction. The allergenic agents (haptens) are typically low-molecular-weight chemicals that are electrophilic. These agents are seldom allergenic alone and must be linked in skin with a carrier protein to form a complete allergen. Some chemicals must be metabolically activated in order to form an allergen, which can occur within the skin as a result of the skin's Phase I and Phase II metabolic activities.

Sensitization occurs when the hapten/carrier protein (antigen) is engulfed by an antigen-presenting cell (e.g., macrophages and Langerhans cells) and the processed antigen is presented to a helper T cell (CD4+). The T cell produces cytokines that activate and cause the proliferation of additional T cells that specifically recognize the antigen. The secretion of cytokines also causes inflammation of the contact area and activation of monocytes into macrophages. The active macrophages are the ultimate effector cells of the reaction. They act to eliminate the foreign antigen and, through secretion of additional chemical mediators, enhance the inflammation of the contact site. Keratinocytes also play a role in the hypersensitivity reaction. They are capable of producing many cytokines and can act as antigen-presenting cells under certain circumstances. After the

TABLE 9.3 Commonly Encountered Contact Allergens

Source	Allergen(s)	Examples
Plants and trees	Rhus	Poison oak and ivy
Metals	Nickel and chromium	Earrings, coins, and watches
Glues and bonding agents	Bisphenol A, formaldehyde, acrylic monomers	Glues, building materials, and paints
Hygiene products and topical medications	Bacitracin, neomycin, benzalkonium chloride, lanolin, benzocaine, and propylene glycol	Creams, shampoos, and topical medications
Antiseptics	Chloramine, glutaraldehyde, thimerosal, and mercurial	Betadine
Leather	Formaldehyde and glutaraldehyde	
Rubber products	Hydroquinone, diphenylguanidine, and *p*-phenylenediamine	Rubber gloves and boots

sensitization process, subsequent exposure to the allergenic chemical triggers the same cascade of events. However, the prior sensitization reaction resulting in a population of T cells specific for the antigen allows the cascade of events to proceed much faster.

Table 9.3 lists some common agents that induce allergenic dermatitis. The actual number of potential allergenic agents is large. Individual sensitivity to a particular allergen varies greatly and is dependent on many factors, as discussed for irritant contact dermatitis. The genetic makeup of the person probably plays a major role in determining whether a response occurs. This is similar to the variability noticed among individuals for their sensitivity to IgE-mediated allergies, such as hay fever, in which some people respond while others do not.

Systemic Contact Dermatitis

A unique situation arises when a contact allergen is ingested or enters the systemic circulation. The most serious effects include generalized skin eruption, headache, malaise, and arthralgia. Flaring of a previous contact dermatitis, vesicular hand eruptions, and eczema in flexor areas of the body may be less dramatic disturbances. Systemic exposure can trigger a delayed type IV hypersensitivity reaction with subsequent deposition of immunoglobulins and complement in the skin, which are potent inducers of the secondary inflammatory response. Therefore, systemic exposure to a contact allergen may induce a widespread delayed type IV hypersensitivity reaction that is not localized to one area of the body.

Diagnostic Patch Testing Diagnostic patch testing is used to try to determine to which agent a person with suspected allergic contact dermatitis may be sensitive. Unfortunately, the test is usually standardized to agents that are the most frequent causes of allergic contact dermatitis. As such, identifying sensitivity to an agent unique to a given occupation or environmental exposure situation may be difficult. Patch testing should be performed by physicians trained and experienced in the technique, its pitfalls, and the subtleties of interpretation. If a compound is identified as allergenic, the sensitive individual can attempt to avoid exposure to that agent. The distribution of the allergic response on the body can also provide clues as to what the allergenic compound is. For example, linear stripes may indicate plant-induced dermatitis, while a rash on the lower abdomen may indicate an allergy to a nickel-containing pants button. A variety of treatments are used to help alleviate contact dermatitis. The best treatment, however, is avoidance of the allergen or irritant. Baths, wet compresses, and highly potent topical and/or oral corticosteroids are used in various combinations to treat contact dermatitis. Fortunately, resolution occurs without therapy when the allergen exposure ceases.

Predictive Testing Testing in guinea pigs, mice, and humans can induce skin sensitization and elicit allergic contact hypersensitivity in a dose-dependent manner and provide estimates of hazard, risks, and relative potency. The threshold for inducing sensitization in a previously naive animal is the concentration of allergen necessary to provoke a cutaneous immune response of sufficient degree to result in sensitization. The threshold for elicitation is the lowest challenge concentration of the previously induced allergen that is required to elicit a detectable cutaneous immune response. Guinea pig sensitization tests include the Draize test, the open epicutaneous test, the Buehler assay, Freund's complete adjuvant test, the optimization test, the split adjuvant, and guinea pig maximization. The local lymph node assay is a novel predictive method for identifying skin sensitization chemicals by measuring the ability of topically applied allergens to induce proliferative responses by draining lymph node cells in mice. Human sensitization tests include the complete Schwartz-Peck test, the Shelanksi–Shelanksi test, the repeat insult patch tests, the Draize test, the Griffith–Voss–Stotts test, the modified Draize test, and human maximization.

Photosensitivity (Photoirritation and Photoallergic Contact Dermatitis)

Photosensitivity is an abnormal sensitivity to UV and visible light and can be caused by endogenous and exogenous factors. Wavelengths outside the UV and visible light ranges are seldom involved, since the earth's atmosphere significantly

filters those wavelengths or they are not sufficiently energetic to cause skin damage. In order for any form of electromagnetic radiation to produce an effect, it must first be absorbed. Chromophores, epidermal thickness, and water content all affect the ability of light to penetrate the skin, and those parameters vary from region to region on the body. Melanin is the most significant chromophore, since it can absorb a wide range of radiation from ultraviolet B (UVB) (290–320 nm) through the visible spectrum.

Exposure to intense sunlight causes erythema (redness or sunburn) due to vasodilation of the exposed areas. Inflammatory mediators may be released at these areas and have been implicated in the systemic symptoms of sunburn such as fever, chills, and malaise. UVB is the most important radiation band in causing erythema. Sunlight has up to 100-fold greater ultraviolet A (UVA) (320–400 nm), but UVA is 1000 times less potent than UVB in causing erythema. UVB exposure causes darkening of the skin through enhanced melanin production or through oxidation of melanin. Oxidation of melanin occurs immediately, but offers no additional protection against sun damage. Enhanced melanin production is noticeable within 3 days of exposure. UV exposure also enhances thickening of the skin, primarily in the stratum corneum, which further retards subsequent UV absorption. Chronic exposure to UV light can induce a number of skin changes such as freckling, wrinkling, and precancerous and malignant skin lesions. UV light is not the only type of radiation that can induce skin changes. Depending on the dose delivered, ionizing radiation can cause acute changes such as redness, blistering, swelling, ulceration, and pain. Following a latent period or chronic exposure, epidermal thickening, freckling, nonhealing ulcerations, and malignancies may occur.

Photoirritation Photoirritation (phototoxicity), a form of irritant dermatitis, results from topical or systemic exposure to exogenous chemicals. The symptoms are very similar to severe sunburn and include reddening and blistering of the skin. Chronic exposure can result in hyperpigmentation and thickening of the affected areas. Unlike sunburn, phototoxicity often results from exposure to the UVA band, but the UVB band is rarely involved. Phototoxic chemicals are protoxicants (i.e., they are not toxic in their native form) and must be activated by UV light to a toxic form. Phototoxic chemicals readily absorb UV light and become excited to a higher energy state. Once the excited chemical returns to the ground state, it releases its energy, which can lead to production of reactive oxygen species and other reactive products that damage cellular components and macromolecules, ultimately causing cell death. The resulting damage is similar to that caused by irritant chemicals that cause cell death. Protoxicant-induced cell death triggers an inflammatory response that produces the clinical signs of phototoxicity. Dyes (e.g., eosin, acridine orange), polycyclic aromatic hydrocarbons (PAHs; e.g., anthracene, fluoranthene), tetracyclines, sulfonamides, and furocoumarins (e.g., trimethoxypsoralen, 8-methoxypsoralen) are commonly encountered phototoxic drugs and chemicals.

Photoallergic Contact Dermatitis Photoallergic contact dermatitis (photoallergy) is similar to contact allergic dermatitis and is a delayed type IV hypersensitivity reaction. The difference between an allergenic chemical and a photoallergenic chemical is that the photoallergenic chemical must be activated by exposure to light—most often UVA. Once activated, the photoallergen complexes with cellular protein to form a complete allergen that triggers the delayed type IV hypersensitivity reaction. Because it is a hypersensitivity reaction, previous exposure to the phototoxic chemical is required for a response. Subsequent topical or systemic exposure to the photoallergen may induce the hypersensitivity reaction, which has clinical manifestations similar to allergic contact dermatitis (see the subsection on allergic contact dermatitis).

Phototesting Screening tests for evaluating phototoxic potential should begin with examination of the test chemical under UV light. Fluorescence under UV light examination suggests that the chemical may be active and may require further testing. Additional testing can be performed with *in vitro* tests. After familiarization and performance of animal tests, human testing can be undertaken, which is similar to the patch testing used for regular allergens, but the potential allergens are tested in duplicate. One set of patches is removed during the test and irradiated with UV light. By comparing duplicate samples, the physician can determine whether the compound is allergenic and is also a photoallergen.

Contact Urticaria Syndrome (Nonimmunological [NICU] and Immunological [ICU])

Contact urticaria syndrome (CUS) encompasses immediate contact inflammatory reactions that appear within minutes after contact with a substance. Contact urticaria refers to a wheal and flare reaction and/or eczema following external contact with a substance, usually appearing within 30 min and clearing completely within hours, typically without residual signs. CUS may be produced by nonimmunological (irritant) or immunological (allergic) mechanisms. These conditions should be distinguished from allergic contact dermatitis, which is a delayed type of induced sensitivity that develops hours to days after contact with the offending agent and manifests as varying degrees of erythema, edema, and vesiculation.

CUS can be described in two broad categories: nonimmunological contact urticaria (NICU) and immunological contact urticaria (ICU). NICU, the most common form of the disease, is due to vasogenic mediators and does not require presensitization of the patient's immune system to an allergen, whereas ICU does. ICU reflects a type I

hypersensitivity reaction, mediated by allergen-specific immunoglobulin E in a previously sensitized individual. Patients with atopic dermatitis are predisposed to ICU, but no other predisposing factors have been described. ICU may spread beyond the initial site of contact and progress to generalized urticaria or anaphylactic shock, which may be seen in ICU secondary to latex protein.

Protein contact dermatitis (PCD) represents a subset of immunological mechanisms in which repeated exposure to the urticariant causes the urticarial lesions to transform into dermatitis (eczema). PCD is caused by recurrent exposure to high-molecular-weight proteins that eventually results in an eczematous dermatitis. It has been reported with increasing frequency, most commonly as occupational hand dermatitis in food handlers. Proteins that induce PCD are classified into four groups: (i) fruits, vegetables, spices, plants; (ii) animal proteins; (iii) grains; and (iv) enzymes. An association between atopy and PCD occurs in approximately 50% of affected patients. PCD typically affects the hands, especially the fingertips, and may extend proximally to include the wrist and arms.

Manifestations of ICU depend on the dose and exposure route of the responsible substance. Skin lesions include wheal and flare reactions as well as symptoms such as itching, tingling, or burning. Wheals appear as the unique symptom in contact urticaria and eczema can be the only symptom in PCD. ICU can be divided into four stages of progression: (i) localized reaction (redness and swelling) with nonspecific symptoms (itching, tingling, burning); (ii) generalized reaction; (iii) extracutaneous symptoms (respiratory or gastrointestinal tract); and (iv) anaphylactic shock and, rarely, death.

Diagnosis is based on a complete medical history and skin testing with suspected substances. Skin testing may not be necessary if the patient is able to correlate symptoms to exposure to a specific substance. Details with regard to the patient's employment and activities at the onset of symptoms may assist in targeting a specific substance. If skin testing is warranted, it should be gradated and carefully monitored to manage potential anaphylactic reactions. *In vitro* techniques are available for only a few allergens, including latex.

The simplest cutaneous provocation test for ICU, NICU, and immediate contact dermatitis is the open test, in which the test substance is applied and gently rubbed on slightly affected skin or on a normal looking area of the skin, either on the upper back or on the flexor side of the upper arm. A positive reaction is characterized by a wheal and flare response. If the open test results are negative, the next diagnostic method of choice for immediate contact reactions is the prick test. The test substance is applied to the flexor side of the forearm and pierced into the skin using a lancet. As only a small amount of test substance is introduced into the skin, there is a low risk of anaphylaxis. If internal organs have been involved in prior urticarial episodes, it is imperative to begin ICU testing with diluted allergen concentrations and to use serial dilutions to minimize allergen exposure. Open and prick tests are also commonly used when testing for PCD. Patch test results are usually negative unless there is a concomitant allergic contact dermatitis.

Toxic Epidermal Necrolysis

Toxic epidermal necrolysis (TEN) is one of the most immediate life-threatening skin diseases caused by chemicals or drugs—typically by oral exposure. Mortality is usually 25–30%, but can be as high as 75% in the elderly. Luckily, the incidence of TEN is low, with approximately one person per million per year becoming affected. The disease is characterized by sudden onset of large, red, tender areas involving a large percentage of the total body surface area. As the disease progresses, necrosis of the epidermis with widespread detachment occurs at the affected area. Once the epidermis is lost, only the dermis remains and this severely compromises the ability of the skin to regulate temperature, fluid, and electrolyte homeostasis. Since the epidermis is lost, the remaining dermis has little resistance to chemicals entering the systemic circulation and to infection from microorganisms.

The mechanism of drug induction of the disease has remained elusive. Recent evidence has implicated immunological and metabolic mechanisms, but they are far from conclusive. TEN has been associated with graft–host disease, and, even though it is a controversial area, TEN is believed to be part of the same spectrum of disease as Stevens–Johnson syndrome, which is another serious reaction to drugs and infection.

Acneiform Dermatoses

Acne is a disfiguring ailment, but fairly innocuous in terms of producing long-lasting damage to the skin. In the workplace, the most common causes of acne are petroleum, coal tar, and cutting oil products. They are termed comedogenic because they induce the characteristic comedo, which is either open (blackhead) or closed (whitehead). The black color of open comedones is due to pigmentary changes resulting in accumulation of melanin. The comedogenic agents produce biochemical and physiological alterations in the hair follicle and cell structure that cause accumulation of compacted keratinocytes in the hair follicles and sebaceous glands. The keratinocytes clog the hair follicles and sebaceous glands and become bathed in sebum (released from the sebaceous glands).

Halogenated chemicals—especially polyhalogenated naphthalenes, biphenyls, dibenzofurans, and contaminants of herbicides such as polychlorophenol and dichloroaniline—cause a disfiguring and recalcitrant form of acne called *chloracne*. Chloracne is typically characterized by the presence of many comedones and straw-colored cysts behind

the ears, around the eyes, and on the shoulders, back, and genitalia. Other symptoms that may or may not occur include conjunctivitis and eye discharge due to hypersecretion of the Meibomian glands around the eyelids, hyperpigmentation, and increased hair in atypical locations. Since chloracne is a persistent disease, the best method of treatment is to prevent exposure to the halogenated chemicals. This could involve putting up splash guards and other devices to prevent the chemicals from coming into contact with the skin along with changing chemical-soaked clothing frequently. Systemic drugs may also produce acneiform eruptions.

Pigment Disturbances

Some chemicals can either increase or decrease pigmentation. These chemicals often cause hyperpigmentation (darkening of the skin) by enhancing the production of melanin or by causing deposition of endogenous or exogenous pigment in the upper epidermis. Hypopigmentation (loss of pigment from the skin) can be caused by decreased melanin production and/or loss, melanocyte damage, or vascular abnormalities. Some common inducers of hyperpigmentation are coal tar compounds, metals (e.g., mercury, lead, arsenic), petroleum oils, and a variety of drugs. Phenols and catechols are potent depigmentors that act by killing melanocytes.

Skin Cancer

Skin cancer is the most common neoplasm in humans, with half a million new cases occurring per year in the United States. Even though exposure to UV light is the primary cause of skin cancer, chemicals can also induce malignancies. UV light and carcinogenic agents induce alterations in epidermal cell DNA. These alterations can lead to permanent mutations in critical genes that cause uncontrolled proliferation of the affected cells, ultimately leading to a cancerous lesion. UVB rays are the most potent inducers of DNA damage and work by inducing pyrimidine dimers. In addition to inducing DNA damage, UV light also has an immunosuppressive effect that may reduce the surveillance and elimination of cancerous cells by the immune system. Since UVB light is the most potent inducer of DNA damage, utilization of a sunscreen that blocks UVB radiation is critical in preventing skin cancer along with the other skin effects associated with UV light exposure. Ionizing radiation is also a potent inducer of skin cancer. Fortunately, ionizing radiation is no longer used for treatment of skin ailments such as acne and psoriasis, as was done in the recent past.

The best characterized chemical inducers of skin cancer are the PAHs. In the 1700s, scrotal cancer was found to be prevalent among chimney sweeps in England. The compounds that induced the cancer were later determined to be PAHs present in high concentrations in coal tar, creosote, pitch, and soot. PAHs must be bioactivated within the skin, often to a reactive epoxide, by cytochrome P450 metabolism in order to cause DNA damage. The epoxides are electrophilic and can form DNA adducts that may produce gene mutations. Other carcinogenic agents may cause DNA damage through different mechanisms, but the ultimate lesion is a gene mutation that leads to a cancerous lesion.

9.4 SUMMARY

Toxicity of the skin can occur after exposure to many different substances that cause injury through a variety of mechanisms. This chapter covered the major problems caused by chemicals encountered at home and at work, but a variety of other skin diseases, including the ones mentioned in this chapter, can occur in reaction to systemically administered drugs. Whether a chemical can produce an effect after it comes in contact with the skin depends on many factors, including genetic makeup, health status, and efficiency of the skin's barrier function. The following are some important points about skin toxicity.

- The most common skin diseases are irritant and allergic contact dermatitis, with allergic contact dermatitis usually being more severe.
- A person must first be sensitized to a chemical before allergic contact dermatitis can occur. Since allergic contact dermatitis is an immune reaction, minute quantities of allergen can trigger the reaction, which makes management of future flare-ups difficult.
- Urticaria may or may not occur through immunity-related mechanisms. The ultimate trigger of the hives associated with urticaria is due to the release of histamine and vasoactive agents from mast cells that are activated after chemical exposure.
- Phototoxicity and photoallergy are similar to irritant and allergic contact dermatitis, respectively. The difference is that the phototoxicant or photoallergen must be activated by exposure to UV light.
- Skin cancer is the most prevalent form of cancer. Its main cause is exposure to UV light, but many chemicals can induce cancerous lesions too, such as PAHs and arsenic.

SUGGESTED READING

Bradley T, Brown RE, Kucan JO, Smoot EC, Hussmann J. Toxic epidermal necrolysis: a review and report of the successful use of biobrane for early wound coverage. Ann Plastic Surg 1995;35:124–132.

Bronaugh RL, Maibach HI. *Percutaneous Absorption: Drugs, Cosmetics, Mechanisms, Methods.* 4th ed. New York: Informa Healthcare; 2005.

Chew AL, Maibach HI. *Irritant Dermatitis*. Berlin: Springer; 2005.

Klaassen CD, editor. *Casarett and Doull's Toxicology: The Basic Science of Poisons.* 7th ed. New York: McGraw-Hill Publishing Company; 2008.

Gattu S, Maibach HI. Modest but increased penetration through damaged skin: an overview of the in vivo human model. Skin Pharmacol Physiol 2010;24:2–9.

Goldstein SM, Wintroub BU. *Adverse Cutaneous Reactions to Medication*. Baltimore: Williams & Wilkins; 1996.

Grandjean P. *Skin Penetration: Hazardous Chemicals at Work*. New York: Taylor & Francis; 1990.

Haschek WM, Rousseaux CG. *Handbook of Toxicologic Pathology*. San Diego: Academic Press; 1991.

Hayes BB, Patrick E, Maibach HI. Dermatotoxicology,. In: Hayes AW, editor. *Principles and Methods of Toxicology.* 5th ed. New York: Informa Healthcare; 2007. pp. 1360–1405.

Hogan DJ. Review of contact dermatitis for non-dermatologists. J Florida Med Assoc 1990;77:663–666.

Lachapelle JM, Maibach HI. *Patch Testing and Prick Testing.* 2nd ed. Berlin: Springer; 2009.

Rice RH, Cohen DE. Toxic responses of the skin. In: Klaassen CD, editor. *Casaretts and Doull's Toxicology: The Basic Science of Poisons.* 5th ed. New York: McGraw-Hill; 1996. pp. 529–546.

Rietschel RL. Dermatotoxicity: toxic effects in the skin. In: Williams PL, Burson JL, editors. *Industrial Toxicology: Safety and Health Applications in the Workplace*. New York: Van Nostrand-Reinhold; 1985. pp. 138–161.

Scheuplein RJ, Blank IH. Permeability of the skin. Physiol Rev 1971;51:702–747.

Taylor JS, Praditsuwan P. Latex allergy: review of 44 cases including outcome and frequent association with allergic hand eczema. Arch Dermatol 1996;132:265–271.

Wang RGM, Knaak JB, Maibach HI. *Health Risk Assessment: Dermal and Inhalation Exposure and Absorption of Toxicants*. Ann Arbor (MI): CRC Press; 1993.

Wilhelm KP, Zhai H, Maibach HI, editors. *Dermatotoxicology.* 8th ed. London: Informa Healthcare; 2012.

Zug KA, McKay M. Eczematous dermatitis: a practical review. Am Family Phys 1996;54:1243–1250.

10

PULMONOTOXICITY: TOXIC EFFECTS IN THE LUNG

CUIQING LIU AND QINGHUA SUN

The lung is the organ primarily where the exchange of oxygen (O_2) and carbon dioxide (CO_2) between the atmosphere and vascular system takes place. It is also an entry and target organ for many inhaled toxins and toxicants, rendering the lung as one of the most vulnerable target organs resulting from exposure to commonly encountered toxicants. The chapter will mainly discuss:

- Anatomy and physiology of the respiratory system
- General principles of inhalation toxicology
  - Toxic inhalants
  - Gas toxicity
  - Particle size, deposition, and clearance
- Acute responses of the respiratory system to injury
- Chronic responses of the respiratory system to injury
- Evaluation of toxic damage in the respiratory system

10.1 ANATOMY AND PHYSIOLOGY OF RESPIRATORY SYSTEM

The primary function of the respiratory system is to obtain O_2 for use by body's cells and eliminate CO_2 that cells produce. During this exchange the respiratory system also assumes an entry for "toxic" substances suspended in the inhaled air. Inhalant uptake is of importance for both occupational and environment exposures. It is essential to understand the anatomy and physiology of the respiratory system, which primarily include the respiratory tract and the lungs.

Respiratory Tract

The respiratory tract is the path that air takes when inhaled through the nose to the lungs, and can be divided into three main compartments, including the extrathoracic, tracheobronchial, and alveolar compartments.

The extrathoracic compartment comprises the nose, mouth, nasopharynx, oropharynx, and larynx. Air enters through the nostrils of the nose and is partially filtered by the nose hairs, and it then flows into the nasal cavity. The nasal cavity is lined with epithelial tissue, containing blood vessels that help warm the air, and secrete mucous that further filters the air. The endothelial lining of the nasal cavity also contains tiny hair-like projections that are called cilia. The cilia serve to transport dust and other foreign particles, trapped in mucous, to the back of the nasal cavity and the pharynx. In addition, the sinuses (including maxillary sinuses, frontal sinuses, ethmoid sinuses, and sphenoid sinuses) are lined with soft, pink tissue called mucosa. In order to view these sinuses from different angles, Figure 10.1 shows a fontal view of the skull, while Figure 10.2 represents a sagittal view. Since the sinuses are connected to the nasopharyngeal airways through a number of small openings, inhaled air also enters the sinuses. When airborne irritants are inhaled the surfaces of sinus mucosa can be the first tissue to be irritated; this irritation may induce sinusitis and possibly lead subsequently to the growth of bacteria. Many factors can contribute to sinusitis, in addition to or in conjunction with inhaled toxins, such as allergic hypersensitivity, individual characteristics of the sinuses in each person, and climatic conditions. After passing through the nasal cavity, the air flows down the pharynx to the larynx.

Principles of Toxicology: Environmental and Industrial Applications, Third Edition. Edited by Stephen M. Roberts, Robert C. James and Phillip L. Williams.

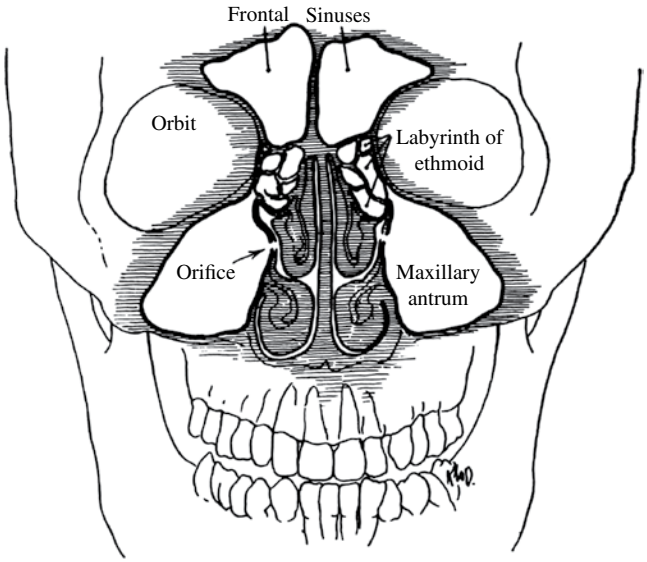


FIGURE 10.1 Frontal view of the skull, showing frontal, maxillary, and ethmoid sinuses. *Source:* Reproduced with permission from Fenn and Rahn (1964).

The tracheobronchial compartment includes the trachea, bronchi, and terminal bronchioles (Figure 10.3). The trachea, a tube surrounded with cartilaginous rings, carries air from the throat into the lung. It ranges from 20 to 25 mm in diameter and 10 to 16 cm in length in humans. The inner membrane of the trachea is also covered with cilia that catch particles in the dusts that can then be removed through coughing. The trachea connects the nasopharyngeal region with the two main bronchi, one entering left and the other entering right. The left bronchus is narrower, longer, and more horizontal than the right. Irregular rings of cartilage surround the bronchus, whose walls also consist of smooth muscle. Once inside the lung, the bronchi now divide up to 16–20 times and become the smaller airways or bronchioles. The bronchi themselves do not allow for the absorption of O_2 or CO_2 across their surfaces but are merely conducting airway tubes. The inner lining of the bronchi is covered by epithelial cells, which include ciliated cells, mucus-secreting goblet cells, serous cells, basal cells, and nonciliated bronchiolar epithelial cells. Here the synchronization of the ciliary beat is essential to the unidirectional movement of the

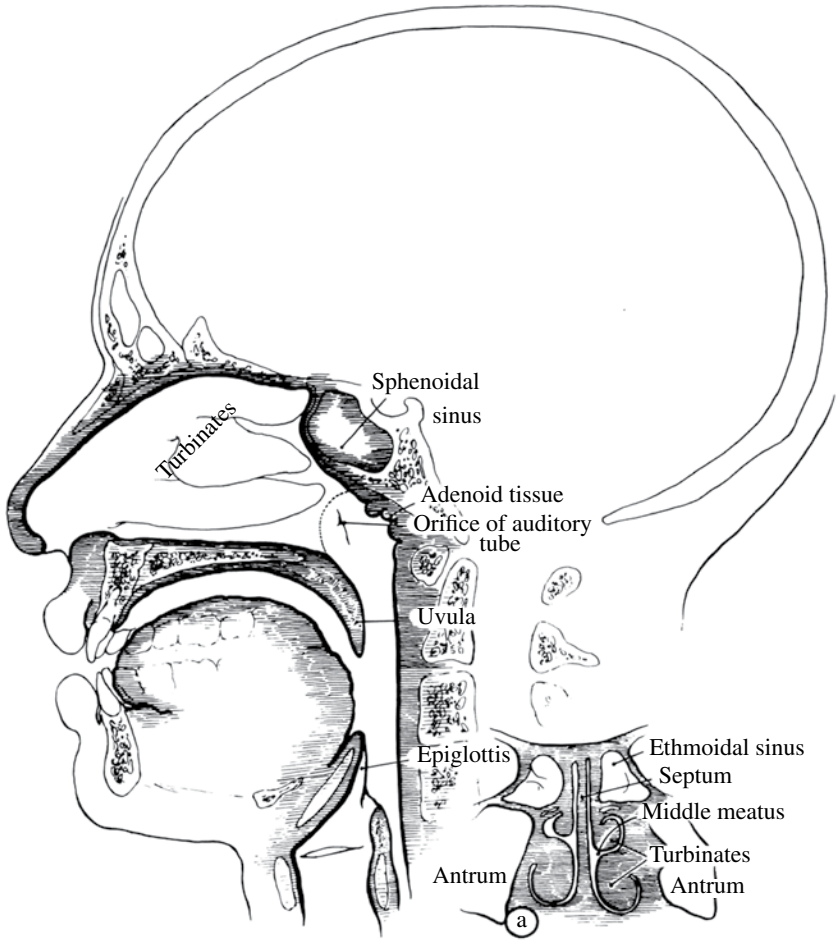
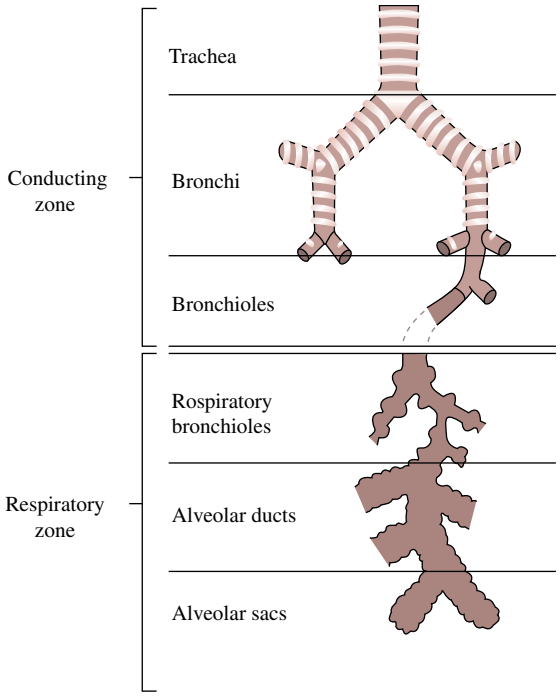


FIGURE 10.2 Saggital view of the skull, showing nasal turbinates and sphenoid sinuses. *Source:* Reproduced with permission from Fenn and Rahn (1964).



Number	Cilia	Number of subdivisions
1	Yes	1
2 4 8	Yes	2–4
—	Yes	4–12
—	Some	13–18
—	No	19–22
6×10^8	No	2–3

FIGURE 10.3 Schematic representation of the subdivisions of the conducting airways and terminal respiratory units.

mucus, which is transported by ciliary activity back toward the pharynx. It is by this mechanism that deposed particles can be removed from the tracheobronchial compartment and leave the body.

Lung

The lung or alveolar compartment consists of the respiratory bronchioles, alveolar ducts, alveolar sacs, and alveoli, with a dramatic increase in the cumulative surface. In the bronchi proximal to the lung, very small air sacs or alveoli begin to appear. The bronchi in this region are known as respiratory bronchioles. The respiratory bronchioles lead to more than 10 million alveolar ducts, which in turn lead to more than 500 million alveoli, the termination of bronchioles. It is in these alveoli that gas exchange between the inhaled air and the blood circulatory system occurs.

The surface of the alveoli is covered by two epithelial cell types, alveolar type I and alveolar type II cells. The alveolar type I cells are thin and cover about 90% of the alveolar surface. Alveolar type II cells are cuboidal and are more numerous, but cover only about 10% of the alveolar surface. These cells can secrete surfactants that line the inner surface of alveoli, which reduces the surface tension of the alveoli and is therefore very important for pulmonary mechanic movement and pulmonary function.

The alveoli comprise approximately 80–90% of the total parenchymal lung volume with a total area for gas exchange of about 100 m^2. In each alveolus, only a thin wall separates the blood in the capillary vessels from the inhaled air in the alveolus. This thin wall or respiratory membrane is a combination of the capillary endothelium, a basement membrane adjacent to the capillary, the space between the capillary and the alveolus (known as the interstitial space), a basement membrane adjacent to the alveolus, and the alveolar epithelium. Although the respiratory membrane consists of five layers, its thickness is only 0.4–2.5 µm. Therefore, CO_2 and O_2 readily cross this membrane by simple diffusion.

10.2 GENERAL PRINCIPLES OF INHALATION TOXICOLOGY

As stated earlier, the large surface area of the lung produced by the alveolar ducts and alveoli is in continuous contact with the inhaled environmental atmosphere and the toxic inhalants present in it, rending the lungs as sites of action for adverse effects for volatile and particulate bound chemicals that might be present in the air. The main toxic inhalants are gas and airborne particles. Toxic gases refer to airborne chemicals that normally occur in the gaseous state and may have contact with the lungs or enter the body mainly by inhalation. Some well-known toxic gases include, but are not limited to, nitrogen dioxide (NO_2), ozone (O_3), sulfur dioxide (SO_2), hydrogen sulfide (H_2S), and chlorine (Cl_2), and so on.

The pattern of toxicity induced by these gases is in part determined by the sites of gas deposition in the respiratory tract. Water solubility is one critical factor in determining

how deeply a given gas penetrates into the lung. High-soluble gases such as SO_2 typically do not penetrate farther than the upper respiratory tract, and it is therefore relatively free of systemic toxicity. This is especially true in obligatory nose breathing species like the rat. In contrast, O_3 and NO_2 are relatively insoluble gases and so are not easily cleared by the upper respiratory tract and can penetrate deeply into the lung reaching the smallest airways and the alveoli where they are capable of eliciting toxic responses. In addition, very insoluble gases such as H_2S efficiently pass through the respiratory tract, are taken up by the pulmonary blood supply, and are distributed throughout the body where toxicity and injury are induced.

In addition, the condition of the alveolar–capillary membrane is important to gas diffusion. Once gas enters the alveolar space of the lung, it can cross the relatively permeable alveolar–capillary membrane complex and enter the pulmonary blood circulation. But poor health conditions in a patient might lead to the engorgement of the interstitial space with fluid, which would impair the diffusion of toxic chemicals across the alveolar–capillary membrane. Although preventing the free exchange of O_2 and CO_2 it may protect the affected individual from the toxic effects of the inhaled toxins that produce systemic toxicity once absorbed into the bloodstream.

Gases

Nitrogen Oxides (NO_x) NO_x comprises a group of gaseous chemicals that include nitric oxide (NO), NO_2, nitrous oxide (N_2O), dinitrogen dioxide (N_2O_2), nitrogen trioxide (NO_3), dinitrogen trioxide (N_2O_3), dinitrogen tetraoxide (N_2O_4), and dinitrogen pentoxide (N_2O_5). Excluding N_2O, the various NO_x are interconvertible and many of them actually coexist in the atmosphere. However, from either an occupational, environmental exposure, or health perspective, the materials of most concern are NO_2 and NO.

A large amount of inhaled NO_2 is removed from the respiratory tract, and absorption of up to 90% of the amount inhaled occurs in both human and laboratory animals. Absorption occurs along the entire tracheobronchial tree and within the respiratory (alveolar) region, but the major dose to tissue is delivered at the junction between the conducting and respiratory airways. Beyond this zone, a dramatic falloff in dose delivered to tissue occurs due to the rapid increase in lung surface area. Acute, high-level exposure to NO_2, which may occur accidentally or in occupational settings, has clear health consequences in humans by producing emphysema-like changes and alterations in antimicrobial defenses. However, the extent to which adverse health effects may occur with long-term exposure to lower levels more relevant to ambient outdoor or indoor environments is yet to be resolved.

Although the mechanism(s) underlying NO_2-induced respiratory toxicology are not well known, data suggest the following: (i) NO_2 exposure affects the nonspecific defense function clearance. Both tracheobronchial (due to mucociliary transport) and alveolar (due to macrophage activity) clearance studies support a graded response, whereby low NO_2 levels accelerate and high levels retard clearance; most effects of NO_2 on clearance seem to begin at higher than ambient levels. (ii) NO_2 impairs resistance to infectious agents (bacteria and viruses) in animals exposed to levels as low as 0.5 ppm for 3 months. Data suggest that short-term repeated exposures may result in a reduction in counts of certain lymphocytes in the lungs or spleen, or a depression in antibody responsiveness to particular antigens. (iii) NO_2 may enhance immune responsiveness by increasing the severity of pulmonary inflammation in sensitized lungs, and may play some role in the exacerbation of immune-mediated respiratory disease. (iv) Another explanation is that NO_2 exposure may produce morphological alterations in the respiratory tract, especially in the centriacinar region, where the conducting and gas exchange airways meet. In the alveolar region, NO_2 results in hypertrophy and hyperplasia of type I cells, followed by death and desquamation of these cells and proliferation of and replacement by type II cells, generating thickened the air–blood barrier. Bronchiolar response is characterized by hypertrophy and hyperplasia of epithelial cells, loss of secretory granules and surface protrusions in Clara cells, and loss of ciliated cells or cilia.

Little is known about NO absorption and even less is known about its subsequent distribution. Because endogenous NO is involved in numerous physiological processes, the impact of inhaled, exogenous NO, especially at low concentrations, is often difficult to evaluate. In spite of any binding with hemoglobin, anoxia of O_2-sensitive organs does not seem to occur, at least with NO levels not higher than 10 ppm. Additionally, methemoglobin, the product of NO and hemoglobin binding, is easily converted into ferrous hemoglobin by methemoglobin reductase. So, as long as the activity of methemoglobin reductase is maintained, the toxic effects for NO is much mitigated than NO_x.

O_3 O_3 is the primary oxidant of concern in photochemical smog due to its inherent bioreactivity. It is a bluish, explosive, irritating, and highly toxic gas. Because of its poor water solubility, a substantial portion of O_3 penetrates deeply into the lung although some 30% is scrubbed by the nasopharynx of humans. One of the biological actions of O_3 is the reaction with unsaturated fatty acids. The strongly electrophilic nature of O_3 enables it to attack carbon double bonds. In the presence of water, the ozonide undergoes hydrolysis with the formation of a carbonyl species and an intermediate zwitterion that ultimately decomposes into another carbonyl compound along with hydrogen peroxide. The ozonization of these fatty acids is essentially equivalent to lipid peroxidation and occurs primarily on biomembranes. Moreover, sulfhydryl compounds may scavenge O_3 with the formation

of disulfides and sulfonic acids. During the reactions of O_3 with oxidizable compounds, reactive oxygen species (ROS) are produced, such as singlet oxygen, OH radicals, and hydrogen peroxide. These secondary reactive compounds may be responsible for subsequent reactions with amino acids and DNA.

The primary sites of O_3 injury are the epithelium of the nasal cavity, trachea and central acinar region. Ciliated cells in the nose and trachea, type I cells in alveolar ducts, and alveolar macrophages are the primary cell types injured. Classically, response of the respiratory system to O_3 exposure can be characterized in terms of three stages: the initial response, proliferation and repair of the epithelium, and the response to continued exposure. Initial responses include injury and death of ciliated cells in the conducting airways and squamous epithelial cells in the parenchyma. The next stage is characterized by exfoliation of the epithelium and increased exudate. Inflammatory cells first appear in the interstitium and then in the epithelium, before they are present in the exudate. The third stage is marked by proliferation of the epithelium, concurrent with downregulation of intraluminal exudates. Significant numbers of inflammatory cells may still be found migrating through the epithelium at this stage, but within 7 days the acute inflammatory response is typically almost resolved. At this time, epithelial proliferation ceases, epithelium is often hyperplastic, and proliferation of matrix components is in progress. Then if O_3 inhalation ceases, the assumption is that the affected compartments will revert to steady state within 7–10 days, which suggests that exposure during repair modifies this process. The effects of long-term exposure are characterized by continuous hyperplasia and low-grade chronic inflammation with exudative cells, as well as synthesis of collagen in the matrix. Although dose–response relationships to lasting effects on base-line lung function and structure are less firmly established, the magnitude of the response is dependent upon the inhaled dose of O_3 and varies with the time course of exposure or the site of the airway tree.

Current evidence supports a complex interaction between inhalation of environmental O_3 and the pulmonary immune system. The biological response to O_3 is dependent on genes of innate immunity including surface receptors, intracellular signaling molecules, and the production of downstream proinflammatory cytokines. Understanding the mechanisms that regulate the response to O_3 can provide fundamental insight into gene and environment interactions and will broadly impact our understanding of inflammatory lung disease.

SO_2 SO_2 is a colorless, highly water-soluble chemical irritant. It is absorbed quickly and almost entirely in the nose and upper airway during quiet breathing, but can be delivered to the lower airways during exercise. SO_2 may quickly dissolve in the aqueous lining of the respiratory epithelium, forming acidic species that result in irritation of the mucous membranes, throat, and respiratory tract. Additionally, clinical studies have demonstrated that SO_2 combined with exercise can cause significant bronchoconstriction in an exposure–response relationship. The chemical mechanisms underlying the bronchoconstrictor effects of SO_2 were not well defined until now. Human and animal studies on SO_2 have shown that it can affect various aspects of pulmonary defenses, such as mucociliary transport, alveolar clearance of deposited particles, and pulmonary macrophage function.

H_2S

H_2S is a colorless, irritating, and highly toxic gas with a characteristic odor resembling rotten eggs, which is recognized at 0.025–1 ppm. H_2S is primarily a systemic respiratory poison with cutaneous absorption being negligible. The most prevalent respiratory symptoms upon accidental exposure to H_2S are dyspnea, sore throat, cough, and chest pain. Chronic inhalation of H_2S may cause inflammation and dryness of the respiratory tract with rhinitis, pharyngitis, laryngitis, bronchitis, and pneumonia, the last of which may be related to the inhibitory effect of H_2S on alveolar macrophages and their ability to inactivate bacteria. It also increases the respiratory rate by the stimulation of peripheral chemoceptors. In acute intoxications with high concentrations of H_2S (>500 ppm), rapid olfactory paralysis occurs and death may ensue from respiratory failure with consequent asphyxia and cardiac failure.

Other Respiratory Irritants

There are some other strong respiratory irritants that induce pulmonary damage after inhalation. Although the concentrations of respiratory irritants in occupational settings that produce toxic inhalation are generally very high and sustained only in incidents involving confined spaces, some of these same compounds are occasionally released in community exposures such as transportation accidents or pipeline leaks.

Cl_2 is a dense, acrid, pungent, greenish-yellow gas; it is a highly reactive oxidant gas and has higher solubility in water than NO_2 and O_3, two well-known environmental pollutants. Currently potential human exposure to chlorine inhalation occurs in a variety of settings in the workplace, as a result of inadvertent environmental releases, and even in the home due to household cleaning mishaps. When inhaled, Cl_2 dissolves in the epithelial lung lining fluid, a thin layer of fluid covering the apical surfaces of airway and distal lung epithelial cells, and then reacts with biological molecules, such as low-molecular-weight antioxidants. Injury to airway and alveolar epithelium results from the chemical reaction products formed when Cl_2 interacts with water (HCl and HOCl) and/or other reaction products, such as chloramines, that are formed from the reactions of these chlorinating

species with biological molecules. Subsequent reactions may initiate self-propagating reactions and induce the production of inflammatory mediators compounding injury to pulmonary surfactants, ion channels, and components of lung epithelial and airway cells. Controlled human exposure data suggest that some subjects may be more responsive to the effects of chlorine gas; epidemiological data also indicate that certain subpopulations (e.g., smokers) may be at greater risk of adverse outcome after chlorine inhalation. Although these findings are intriguing, additional study is needed to better delineate the risk factors that predispose toward the development of long-term pulmonary sequelae following chlorine gas exposure.

Hydrogen chloride (HCl) is another irritating chemical that is corrosive to any tissue with which it comes into contact. As it is highly soluble in water, following acute inhalation HCl is deposited in the nose and upper respiratory tract, causing irritation and ulceration, coughing, chest tightness, shortness of breathing, and choking. At higher concentrations, HCl may cause tachypnoea, swelling of the throat leading to suffocation, as well as pulmonary edema. Reactive airway dysfunction syndrome, a chemical-induced type of asthma, may also occur. Chronic inhalation exposure to HCl gas or mist may result in decreased pulmonary function, inflammation of the bronchi, and nasal ulceration. Table 10.1 presents the main gases/chemicals and some of their most relevant chemical and physical characteristics.

Particles

Particulates mean a population of particles that remain suspended in the ambient air over time. In addition to the ability of the physical nature of particulate to induce lung injury if accumulated in the lungs, many particulates contain chemical and radionuclide agents that are also deposited in the respiratory tract in the form of solid particles or droplets. These particles are widely originated and present as aerosols, dust, fumes, smokes, mists, and smog. Dusts result from industrial processes such as sandblasting and grinding, and are identical to the compounds from which they originated. However, fume occurs with combustion or sublimation and derived from chemical change in compounds during processes such as welding. Smokes are produced during burning of organic materials, and mists are aerosols composed of water condensing on other particles. Smog is a conglomerate mixture of particles and gases that is prevalent in certain environments such as areas with mountains, plenty of sunlight, and periodic temperature inversions.

Particle Size

The particle size is usually the critical factor both for the region of the respiratory tract in which a particle or an aerosol will be deposited and for the number of particles that will be deposited. Since most inhaled particles are irregular in shape, but not spherical, particles are frequently characterized in terms of equivalent spheres on the basis of equal

TABLE 10.1 Toxic Gases and Characteristics

Toxic Gas	Chemistry	Water Solubility	Site of Action	Exposure Settings	Toxic Effects in Respiratory System
Nitrogen dioxide (NO_2)	Acid, oxidant	low	Entire airway and deep lung, especially the junction between the conducting and respiratory airways	Chemical industry, high-temperature combustion, nitric acid fumes, farms	Pulmonary edema; dyspnea and bronchiolitis obliterans; lung defense dysfunction
Ozone (O_3)	Strong oxidizing agent	Low	Deep lung	Welding on aluminum, water treatment	Airway inflammation and permeability; decreased respiratory function
Sulfur dioxide (SO_2)	Acid	High	Upper airway	Chemical industry, sulfur combustion, smelting	Bronchitis
Chlorine (Cl_2)	Strong oxidizing agent	High	Airway and distal lung	Pulp mills, water disinfection, chemical industry, home cleaning	Bronchitis; persistent airways reactivity, chronic obstructive changes, BO
Hydrogen chloride (HCl)	Acid	high	Upper airway	Chemical industry, end uses	Bronchitis; pulmonary edema; decreased pulmonary function
Hydrogen sulfide (H_2S)	Reducing agent	low	Entire airway and lung	Oilfields, refining, manure pits, sewer gas	Apnea; pulmonary edema; olfactory paralysis; pneumonia

mass, volume, or aerodynamic drag. Mass median aerodynamic diameter (MMAD) represents the diameter of a unit density sphere with the same terminal settling velocity as the particle, regardless of its size, shape, and density. If the number of particles is of primary interest, the count median diameter (CMD) is used to reflect the number of particles. Aerodynamic diameter is the proper measurement of particles that are deposited by impaction and sedimentation. Particle size is typically critical for very small particles, commonly deposited primarily by diffusion. However, it must be kept in mind that the size of particles may change during the course of traveling in the respiratory tract. Since the respiratory tract is highly humidified, materials that are hygroscopic, such as sodium chloride, sulfuric acid, and glycerol may absorb water and could be expected to increase in size as they descend in the respiratory system.

Particle Deposition

Particles tend to deposit in the lung, which is determined by particle size, respiratory pattern, and regional characteristics of the respiratory tract. For example, very large particles are easy to be blocked out by nose hairs. The deposition of particles occurs primarily by interception, impaction, sedimentation, and diffusion (Brownian movement).

Interception occurs when a particle comes close to a surface so that an edge of the particle contacts the airway surface. This process is important for the deposition of long particles, especially fibers, because it is usual for a fiber to be only 1 μm in diameter but 200 μm in length, leading to enhanced probability of contact with the respiratory tract surfaces. Particles suspended in air tend to continue to travel along the respiratory tract.

In an airstream at an airway bifurcation, a particle may be impacted on the surface. At relatively symmetrical bifurcations, such as the human lung, the deposition rate is likely to be high for particles that move in the center of the airway. In an average adult, most particles larger than 10 μm in aerodynamic diameter are deposited in the nose or oral pharynx and cannot penetrate into the tissues distal to the larynx. Very fine particles (0.01 μm or smaller) are also trapped relatively efficiently in the upper airways by diffusion. Once particles penetrate beyond the upper airways, they are available to be deposited in the bronchial region and the deep-lying airways. Therefore, the alveolar region has significant deposition efficiencies for particles smaller than 5 μm and larger than 0.01 μm.

At the places where the airways are small and the velocity of airflow is low, such as the small bronchi, the bronchioles, and the alveolar spaces in the tracheobronchiolar region, particles in the range of 1–5 μm tend to deposit by sedimentation. This process results from equilibrating between gravitational force and the sum of the buoyancy and the air resistance, and the particle continues to settle with a constant velocity known as the terminal settling velocity.

Diffusion is an important factor in the deposition of submicrometer particles. This Brownian motion increases with decreasing particle size, so deposition is mainly by diffusion in the nose and in other airways and alveoli for particles smaller than 0.5 μm. Besides, the pattern of breathing is an important factor in particle deposition. At the condition of quiet breathing, most of the inhaled particles may be exhaled. However, at higher velocities, a larger volume of air is inhaled than exhaled, so impaction in the large airways and sedimentation/diffusion in the smaller airways and alveoli increase. In a similar manner, breath holding even increases deposition from sedimentation and diffusion further. Another factor that alters particle deposition is to modify the diameter of the conducting airways. For example, cigarette smoking, irritant materials, and partially occluded airways in patients with chronic bronchitis have the potential to increase the deposition of particles by impaction and diffusion in the small airways.

Particle Clearance

The clearance of deposited particles is an important aspect of lung defense. Generally, the clearance mechanism is related to the site of deposition. Respiratory clearance is not synonymous with clearance or systemic clearance from the body in the pharmacokinetic sense. It means that particles or other toxins are removed from the respiratory system whereas the ultimate removal is achieved through the pulmonary blood, the lymphatics, and the gastrointestinal system. The clearance occurs primarily in three compartments of the respiratory system: extrathoracic, tracheobronchial, and alveolar compartments (Figure 10.4). Particles deposited in the extrathoracic compartment, mainly in the

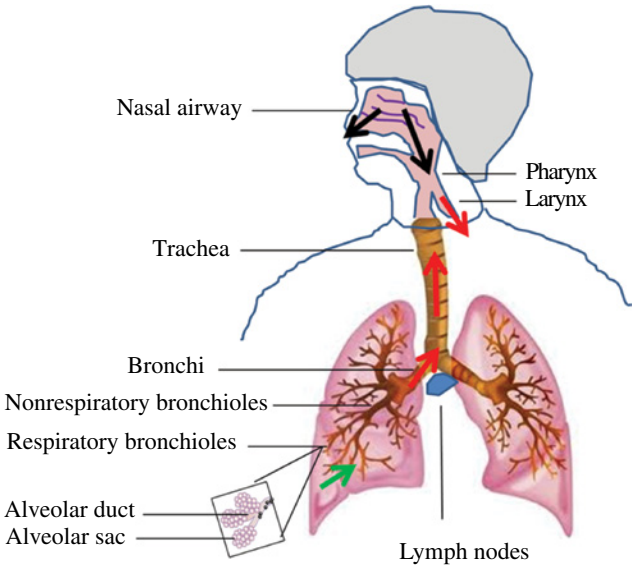


FIGURE 10.4 Principal sites of particle clearance in the lungs.

nose, are cleared depending on their sites of deposition. For example, if the particles are deposited in the anterior portion of the nose, they are removed by extrinsic actions such as blowing the nose. However, the other part of the nose is covered by a mucociliary epithelium that propels mucus toward the glottis, where it can be swallowed. At this region, the solubility of particles in mucus determines the clearance. Insoluble particles generally are cleared at this place because they can be swallowed within an hour of deposition. Soluble particles tend to dissolve and enter the epithelium and/or blood before they can be mechanically removed.

Some special structure, including pseudostratified columnar epithelial cells together with specialized goblet cells and hundreds of cilia, which protrude from the epithelial cells, is present in the respiratory wall of the tracheobronchial compartment. The mucous secreted from the goblet cells is in two layers: thin and watery *sol* in the lower layer and thick and viscous upper layer. Inhaled particles, particle-laden macrophages, and other toxins become trapped on the gel layer. The mucus layer in this region is moved upward by the beating of the underlying cilia, transporting deposited particles or toxins upward to the oropharynx and swallowed into the gastrointestinal tract. This process is termed as mucociliary escalator mechanism and typically completed with 24–48 h. In the alveolar compartment region, macrophages provide an effective defense mechanism against particles or any other offensive agents that are deposited in the lower respiratory tract. In details, when the inhaled agents deposit in the lung, chemotactic factors are released and direct phagocytic cells are recruited to the location of the agents. Then the macrophages engulf the particles and break them down with proteolytic enzymes. Finally, the phagocytized complex is cleared via the mucociliary escalator mechanism or via lymphatics drainage. This action is rapidly processed within minutes of inhalation. However, in certain situations (e.g., unhealthy individuals of long-term tobacco smokers), the insufficient amount of macrophages limits the clearance action. Besides, there are some other ways for particles in the alveolar region to be removed. Particles trapped on the lining layer of the conducting airways by impaction may be cleaned upward via the mucociliary escalator. Some insoluble particles, especially long narrow fibers, may be sequestered in the lung for very long periods, often in macrophages located in the interstitium.

In many instances, some inhaled airborne particles will also readily cross the membrane and may penetrate into the bloodstream. Some potential toxins and toxicants thus enter the blood circulatory system constantly in a low-grade manner. Toxicological insult to the lung as well as various disease states can result in a functional derangement of this membrane system. Exposure to some chemicals may result in an increase in fluid in the interstitial space. If sufficient fluid accumulates, a condition known as pulmonary edema, gas exchange can be hindered to result in severe difficulty in breathing and even in death. As with pulmonary edema, an increase in the thickness of the membrane can deleteriously affect pulmonary gas exchange. Whenever gas exchange is substantially decreased, the amount of O_2 pressure in the circulatory system will also decline, which may lead to a seriously compromised health outcome.

10.3 ACUTE RESPONSES OF THE RESPIRATORY SYSTEM TO INJURY

Airborne agents deposited in the respiratory system can contact with cells lining in the entire respiratory tract from the nostrils to the alveoli. The risk to humans produced by inhalants is highly related with the site of the interaction of toxicants in the respiratory system. For instance, the reactive gas formaldehyde has been demonstrated to develop nasal tumors, evidenced by DNA–protein cross-links formed in nasal tissue. Moreover, contrary to humans, rats, nocturnally active animals, tend to receive a greater dose of inhalants per unit of exposure at night than during the day, indicating that the patterns of animal activity can affect inhalant doses to the lung.

Toxic gases can cause respiratory injury in different ways. Certain gases stimulate nerve endings in the nose, particularly those of the trigeminal nerve. The reaction of humans to toxic gases is to hold the breath or change the breathing patterns in order to avoid further exposure. However, many acidic or alkaline irritants induce cell necrosis and increase permeability of the alveolar walls if the toxic exposure continues. For some inhaled chemicals, there is a latency period between inhalation and occurrence of respiratory injury. For example, the epithelial barrier in the alveolar region begins to leak, flooding the alveoli and producing a delayed pulmonary edema after several hours inhalation of high concentrations of HCl or NO_2. The host defense systems and the function of macrophages are affected by NO_2 exposure, too. O_3 is another toxic gas. At concentrations of 0.25–0.75 ppm, O_3 causes shallow and rapid breathing, a decrease in pulmonary compliance, and other discomforts, such as cough, tightness in the chest, and dryness of the throat. More attention has been paid to the role of inflammatory effects, which has been observed in animal exposure studies. Following O_3 exposure, the bronchoalveolar lavage fluid showed an increase in proinflammatory cytokines, polymorphonuclear leukocytes, together with arachidonic acid descendants, such as the prostaglandins (thromboxane B_2, PGE_2, and PGF_{2a}). Other lavage compounds indicative of inflammation include interleukin (IL)-6, lactate dehydrogenase, albumin, and fibronectin. Inhaled O_3 also causes altered distribution of toll-like receptor (TLR)-4 on alveolar macrophages and enhanced functional response to endotoxin by macrophages. These observations indicate that O_3 exposure increases both

the pulmonary and the systemic biological reactions that may be mediated via priming the innate immune system.

Another factor that is involved in the pathogenesis of respiratory injury is the metabolism of foreign compounds. Enzymes involved in xenobiotic metabolism are highly concentrated in specific cell populations of the respiratory tract and the content of particular cytochrome P-450 isozymes may be much higher in some part of the respiratory system, resulting in more rapid turnover of a particular substrate by P-450. Some isozymes of the cytochrome P-450 complex have been demonstrated to be involved in respiratory injury. For example, cytochrome P-450 1A1 is present in low amounts in normal rat and rabbit lungs but is highly inducible by polycyclic aromatic hydrocarbons (PAH), flavones, and mixtures of polyhalogenated biphenyls, leading to the inference that this P-450 may play a role in the pathogenesis of lung cancer. Whereas cytochrome P-450 2B1 is readily inducible in rat liver by phenobarbital, it is not inducible in lung tissue. In human lung cytochrome P-450 2F1, 4B1, and 3A4, NADPH cytochrome P-450 reductase, epoxide hydrolase, and flavin-containing monoxygenases have been identified. Glutathione-S-transferases and glutathione peroxidase are two other pivotal enzymes involved in lung xenobiotic metabolism.

Burden of Oxidative Stress in Response to Respiratory Injury

Emerging studies have reported the increased activity of free radical scavenging enzymes in the respiratory system of the animals exposed to O_3, NO_2, or other toxicants. By inference, the undue oxidative burden may play a role in respiratory damage. The reactive free radicals and oxygen species can lead to uncontrolled destructive oxidation. It is demonstrated by many studies that superoxide, nitric oxide, peroxynitrate, hydroxyl radicals, hydrogen peroxide, and even singlet oxygen mediate respiratory tissue damage.

The reduction of O_2 to active O_2^- metabolites normally occurs as a by-product of cellular metabolism during both microsomal and mitochondrial electron transfer reactions; NADPH cytochrome P-450 reductase reactions render a great amount of superoxide anion. The cytotoxic oxidant species may mediate or induce the actions of various pneumotoxicants. In addition, the oxygen species can be produced as a by-product of phagocytotic and antimicrobial activities, as neutrophils, monocytes, and macrophages are particularly adept at converting molecular O_2 to reactive O_2 metabolites.

The toxic radicals stimulated by phagocytes are released into surrounding tissues, inducing oxidative damage in pneumotoxic injury. This is supported by phagocyte accumulation in pulmonary microcirculation (pulmonary leukostasis) and parenchyma in most forms of toxic pulmonary edema. A key role of hydrogen peroxide as the mediator of the extracellular cytotoxic mechanism of activated phagocytes has been well documented. Besides, hydrogen peroxide can amplify cell damage by crossing cell membranes and working as a potent intracellular signaling molecule.

Immune Reactivity in Response to Respiratory Injury

The immune system can mount either cellular or humorally mediated responses to airborne microorganisms or various low- and high-molecular-weight antigenic materials. Direct immunological effects occur when the inhaled foreign material sensitizes the respiratory system to further exposure to the same material. Frequently, chemical components of the sensitizing particles or gases are responsible for the allergic response. Low-molecular-weight compounds can act as haptens that combine with native proteins to form a complex that is recognized as an antigen by the immune system. Further exposure to the sensitizing compound can result in an allergic reaction that is characterized by the release of various inflammatory mediators that produce an early and/or a late bronchoconstrictor response. Such a response is observed in sensitized workers exposed to toluene diisocyanate (TDI), a chemical widely used in the manufacture of polyurethane plastics.

The respiratory system can respond to toxic particles or gases by secreting many inflammatory factors. For example, IL-1β, transforming growth factor (TGF)-β, and tumor necrosis factor (TNF)-α have been implicated in the cascade of reactions that are thought to be responsible for the pathogenesis of pulmonary fibrosis. Interleukin family members including IL-1, IL-2, IL-5, IL-8, and IL-13 are essential to the response to epithelial cell injury in the respiratory system. Prostaglandins, especially PGE_2 and leukotrienes, have been focused as intracellular signaling molecules in the respiratory injury. The roles of cell surface adhesion molecules in their interaction with cell matrix components and with control of inflammatory cell migration (particularly neutrophil influx to the respiratory system) have been investigated intensively. For example, diesel exhaust particles (DEP) have been shown to enhance lung injury related to endotoxin through expression of proinflammatory cytokines, chemokines, TLR, and intercellular adhesion molecule-1 by further activating nuclear translocation of p65 subunit of nuclear factor-kappaB (NF-κB) in the lung. According to the lung homogenates assay, there are a great amount of endogenous cytokines and inflammatory mediators, far more than the amount that elicits effects. By inference, these agents should be compartmentalized in a healthy lung to inhibit their potent bioactivity. However, how these processes are regulated in a normal lung and in a damaged lung, the role of inflammation factors in response to the injury, and the detailed mechanisms of lung injury still remain unclear and represent the current focus of much research on mechanisms of respiratory injury by toxic agents.

Alteration of Airway Function

Large airways are surrounded by bronchial smooth muscle, which maintains appropriate airway tone and diameter during respiratory inflation and deflation. The autonomic nervous system is mainly responsible for the bronchial smooth muscle tone. Besides inhalant toxicants that appear to act wholly or partly through an allergic response, bronchoconstriction can also be provoked by irritants including volatile organic compounds, cigarette smoke, ambient air pollutants, and cholinergic drugs such as acetylcholine. The activation of cholinergic receptors on the cell surface triggers an increase in the intracellular concentration of cyclic guanosine monophosphate (cGMP), which in turn facilitates smooth muscle contraction. However, another cyclic nuecleotide called cyclic adenosine monophospate (cAMP) has bronchodilatory activity and can antagonize the action of cGMP. There are a lot of other important mediators of airway smooth muscle tone including histamine, various prostaglandins and leukotrienes, substance P, and nitric oxide. The bronchial smooth muscles of individuals with asthma have higher reactivity than do those of normal subjects by contracting with much less provocation of bronchoconstrictors. Considering bronchoconstriction is one of most typical symptoms in asthma, asthmatic individuals may represent a population that is vulnerable to ambient air pollution, especially O_3, other respiratory irritant gases, and particulates.

Acute Pulmonary Edema

Toxic pulmonary edema represents an acute, exudative phase of lung injury. Many inhaled agents produce sufficient cellular toxicity to cause an increase in the membrane permeability of the respiratory membrane complex in the lung and other airway linings. This leads to an increase in fluid, either in the interstitial space of the respiratory membrane complex or on the surface of the airways or alveolar sacs. This increase in fluid is called edema and its presence limits diffusive transfer of O_2 and CO_2 even in otherwise structurally normal alveoli. When the decrease in gas exchange proceeds sufficiently, the affected individual can die, literally in their own fluids.

Among many agents that result in acute pulmonary edema are the air pollutant gases, such as NO_2 and NO_3. These agents typically exert respiratory toxicity at relatively low levels of exposure in air pollution episodes. However, in industrial exposures, workers may be exposed to considerably higher concentrations, resulting in a high frequency of worker injury. Chlorine is another potent inducer of pulmonary edema, which was used as chemical warfare gas to induce many deaths during World War I. Since chlorine is now the primary chemical used to keep water supplies clean and municipalities use chlorine for their drinking water treatment, its geographic distribution is widespread. There are a number of fatalities from pulmonary edema following chlorine release occurring during its transportation.

The delayed onset of pulmonary edema in most cases of chemical inhalation results in a significant hazard for exposed workers. Usually, the edema fluid is not readily detected by the exposed individual or by clinical examination for at least several hours after the termination of the exposure. Often, when subjects awake with difficulty in breathing, they are already in an advanced stage of pulmonary function decline under a difficult-to-treat condition. Thus, it is necessary for the subjects exposed (or potentially exposed) to agents known to cause pulmonary edema, to be kept at least 24 h following the exposure at a medical facility where they can be closely monitored.

It is worth noting that the initial responses to pulmonary injury may induce subsequent responses (e.g., inflammation mediation, pulmonary fibrosis, and pulmonary edema) that may resolve or lead to permanent changes that produce declines in pulmonary function. After exposure to some toxic chemicals in which the alveolar–capillary surface is denuded (such as alloxan), recovery is unlikely, whereas in situations of more modest injury (such as histamine administration), full recovery is readily achievable. Between these two extremes, there are forms of severe pulmonary injury accompanied by amplified inflammatory damage and/or exaggerated restorative–reparative processes (e.g., after paraquat ingestion). In these severe forms, the extensive interstitial and intra-alveolar inflammatory exudates resolve via fibrogenesis. The accumulation and turnover of inflammatory cells and related immune responses in an edematous lung probablly play a role in eliciting both mitogenic activity and fibrogenic responses.

10.4 CHRONIC RESPONSE OF THE RESPIRATORY SYSTEM TO INJURY

Chronic Obstructive Pulmonary Disease

Chronic obstructive pulmonary disease (COPD) is one of the most common lung diseases, characterized by air flow limitation that is not reversible and is usually progressive. The condition is thought to arise from an abnormal inflammatory response in the lungs to inhaled particles and gases. There are two main forms of COPD: emphysema and chronic bronchitis.

Emphysema is defined as irreversible enlargement of airspaces distal to the terminal bronchioles, accompanied by destruction of the alveolar walls without significant fibrosis. The lung becomes significantly larger and too compliant in the condition of emphysema. The destruction of gas-exchanging tissue results in a distended, hyperinflated lung that no longer effectively exchanges O_2 and CO_2. By far, the leading cause of emphysema is tobacco smoking, whereas

many toxicant particles, for example, coal mine dust, significantly elevated emphysema severity. In an animal study, mice with defects in genes that code for elastin and collagen modifying enzymes develop emphysema, suggesting elastin synthesis may play an important role in the pathogenesis of emphysema.

Proteolytic enzymes are normally counterbalanced by antiproteolytic components, for example, alpha1-antitrypsin (now called alpha1-antiprotease). Both hereditary deficiency and homozygotes for the alpha1-antiprotease gene tend to get emphysema at a very early age. The alpha1-antiprotease is one of the body's main defenses against uncontrolled proteolytic digestion by enzymes, including elastase. Usually, these elastases are inhibited by alpha1-antiprotease that diffuses into the lung from the blood. As an individual, an accumulation of random elastolytic events can cause the emphysematous changes in the lung that is normally associated with aging.

A feature of toxicant-induced emphysema is severe or recurrent inflammation, especially alveolitis with the release of proteolytic enzymes by participating leukocytes. Toxicants cause inflammatory cell influx and thus increase the burden of neutrophil elastase, which can break down lung elastin and thus cause emphysema. Consistent with this, a number of experimental studies in animals instilled intratracheally with pancreatic or neutrophil elastase or with other proteolytic enzymes that can digest elastin.

Chronic bronchitis is defined clinically as the presence of a chronic productive cough on most days for 3 months, in each of two consecutive years, in patients in whom other causes of chronic cough have been excluded. About 90% of all cases occur in smokers, especially in tobacco smokers. Epidemiological studies have shown a correlation between chronic bronchitis and high dust and/or SO_2 concentrations. In coal mining and in foundries, an increased incidence of chronic bronchitis has been reported.

It is generally known that smoking is the leading cause of COPD. Cigarette smoking causes increased risk for respiratory infection or causes inflammation of the lungs, with the migration of inflammatory cells into the lungs and release of enzymes that can destroy the lung's delicate structure. Smoking activates the inflammatory process and reduces the efficacy of defenses against inflammation. Unchecked inflammation, sustained over many years, underlies the development of COPD. In some smokers, this loss is accelerated and eventually lung capacity becomes reduced to a level at which function is affected and the symptoms of COPD occur.

To the infants and children, environmental tobacco smoke exposure has adverse effects on their respiratory health including increased risk for more severe lower respiratory infections, middle ear disease, chronic respiratory symptoms, and asthma, and a reduction in the rate of lung function growth during childhood. There is more limited evidence suggesting that environmental tobacco smoke exposure of the mother reduces birthweight and that child development and behavior are adversely affected by parental smoking. In adults, environmental tobacco smoke exposure has been associated with exacerbation of asthma, reduced lung function, and respiratory symptoms, but these associations have not yet been judged causally.

Tobacco smoke is a complex mixture of gases and condensed tar particles. Its components inhaled and deposited in the lung, with sustained smoking, cause COPD. Regarding its mechanism, it is well known that cigarette components inhaled and deposited in the body impose oxidative stress in the lung by continuous generation of ROS and various inflammatory mediators. Exposure to cigarette smoke causes oxidation of proteins, DNA, and lipids, which may cause direct lung injury or induce a variety of cellular responses through the generation of secondary metabolically reactive species. In addition, cigarette smoke alters remodeling of the extracellular matrix, mucus hypersecretion, plasma exudation/epithelial permeability, cell death, mitochondrial respiration, cell proliferation, maintenance of surfactant and the antiprotease screen, effective alveolar repair response, and immune modulation in the lung. Furthermore, cigarette smoking has been implicated in initiating the lung inflammatory response through the activation of transcription factors such as NF-κB and activator protein-1 (AP-1), signal transduction (activation of MAP kinase pathways), phosphoinositide 3 (PI-3-kinase, PI-3K) and PI-3K-activated serine-threonine kinase, and chromatin modeling (histone acetylation/deacetylation) leading to gene expression of proinflammatory mediators.

Fibrosis

Pulmonary fibrosis may be defined by the presence of an increased amount of newly formed collagen fibers. In the lung that is damaged by toxicants, the response resembles adult or infant respiratory distress syndrome more closely than it resembles chronic interstitial fibrosis. Excess pulmonary collagen is detected not only in the alveolar interstitium but also throughout the centriacinar region, including the alveolar ducts and respiratory.

The formation of collagens is a complex process and the relationship between increased collagen deposition around small airways and lung mechanics is not completely understood either theoretically or empirically. There are emerging studies associating specific macrophage-secreted cytokines with the fibrogenesis. TNF-α, fibronectin, and various growth factors, such as TGF-β, can stimulate fibroblasts to produce collagen. In addition, type I and III collagen are major pulmonary interstitial components, representing about 90% or more of the total lung collagens. In patients with idiopathic pulmonary fibrosis or acute respiratory distress syndrome, type I collagen increases significantly compared to

type III collagen. Since type III collagen is more compliant than is type I, the increase in type I relative to type III collagen may result in a stiffer lung.

Asthma

Asthma is characterized clinically by intermittent symptoms of wheezing, dyspnea, and cough with some degree of persistent airflow obstruction. Typically, airflow limitation is observed due to reversible narrowing of the airways (bronchoconstriction), with edema, mucosal inflammation, and mucus secretion. The clinical literature describes "an epidemic" of childhood asthma, with prevalence rates reportedly as high as 40% in children living in the inner city. This alarming incidence of childhood asthma has stimulated public health concerns, many of which have focused on air pollution (especially ultrafine particulate matter pollution) as a possible cause of the increase in asthma. Epidemiological findings have identified a possible role for O_3 in adult-onset asthma. Sulfate and NO exposure may alter the composition and fate of aeroallergens, thus leading to asthma. Exposure to multiple toxicants, such as diisocyanate, formaldehyde, chromium and nickel salts, and a large variety of naturally occurring dusts, such as flour and grain dusts, and some specific wood dusts, may induce asthma. Air pollution from road traffic (fine particles <2.5 μm in aerodynamic diameter), ultrafine particles (<0.01 μm in aerodynamic diameter), elemental carbon, and NO_2 may exacerbate pulmonary dysfunction in asthmatic children and adults.

The high incidence of childhood asthma has been previously described, indicating that infants and small children may be a particularly sensitive subpopulation to suffer from the adverse health effects caused by air pollutants. It has been recognized that the developing lung is uniquely sensitive to many airborne and blood-borne toxicants. An 8-year period study shows the deficit in pulmonary functional test forced expiratory volume (FEV)1 was associated with exposure to NO_2, acid vapor, particulate matter with an aerodynamic diameter of less than 2.5 μm ($PM_{2.5}$), and elemental carbon. An animal study also provided both anatomical and functional support for the affected lung development by urban particulate matter exposure. Children living in homes with active smokers have been found to suffer from increased infection at the lower respiratory tract (bronchitis and pneumonia), or prone to asthma development. Even infants in a tobacco smoking family may suffer from sudden infant death syndrome (SIDS). It is also reported that fetuses in utero from smoking parents tend to develop SIDS after they are born, and airway narrowing may be one possible mechanism associated with SIDS formation since infants born to tobacco smoking parents had thickened and narrower airways compared to the infants of nonsmoking parents. These early-life alterations may play an important role in the development of lung structure and function in their later lives. Thus, the study of perinatal lung toxicology is a developing field that promises to have considerable impact on human health.

Silicosis

Silicosis refers to a spectrum of pulmonary diseases attributed to the inhalation of various forms of free crystalline silicon dioxide or silica. The gross and microscopic pathology of silicosis are described in detail in a publication on silicosis and silicate by the Disease Committee of the National Institute for Occupational Safety and Health (NIOSH).

Both animal and human data have correlated progression of silicosis with the magnitude of pulmonary ROS production. Siloxyl radicals are generated during cutting, fracturing, or milling quartz, which is silicon dioxide (SiO_2) arranged in a three-dimensional tetrahedral crystal lattice, and produce highly reactive $^{\bullet}OH$. In an aqueous solution, the freshly fractured silica may produce superoxide radicals, singlet oxygen, and these radicals are positively associated with lung injury. Cells involved in the production of these reactive oxidants include alveolar macrophages, polymorphonuclear leukocytes, and alveolar type II epithelial cells. Moreover, freshly fractured silica caused 1.52-fold more oxidant damage (lipid peroxidation), 1.88-fold more lung injury (leak of protein from the alveolar capillaries to the air spaces), and 1.96-fold more inflammation (polymorphnuclear cells obtained by bronchoalveolar lavage) than aged silica.

Data also indicate that oxidant species are involved in DNA damage and loss of cell cycle control, which initiates silica-induced carcinogenesis. Freshly fractured silica has been shown to cause DNA strand breaks in a cellular system, which is blocked by catalase and $^{\bullet}OH$ scavengers (dimethylsulfoxide or sodium benzoate) indicating a role for ROS in this damage. Transcription factors, including NF-κB and AP-1, are associated with lung injury by being involved with silica-induced ROS production in alveolar macrophages and epidermal cells. In addition, cellular, animal, and human data indicate that silica exposure also stimulates NO production in the lung. There is an increasing body of evidence to associate the initiation and progression of silica-induced fibrosis and carcinogenesis with the production of ROS and/or NO. Such an association supports the development of therapeutic strategies to mitigate oxidant generation and oxidant-induced damage and inflammation for management of silica-induced diseases.

Lung Cancer

At the beginning of the twentieth century, lung cancer was an extremely rare disease. According to epidemiological studies, it is now the leading cancer killer that is directly related to tobacco smoking. Average smokers have a 10-fold and heavy smokers a 20-fold increased risk of developing

lung cancer compared with nonsmokers. Long-term exposure to combustion-related fine particulate air pollution is an important environmental risk factor for lung cancer mortality. Even indoor air pollution has been reported to increase the risk of developing lung cancer in nonsmokers. In addition to tobacco smoke inhalation, many industrial chemicals have also been linked to lung cancer in workers and laboratory animal studies. First, inhalation of asbestos fibers and metallic dusts or fumes, such as arsenic, beryllium, cadmium, chromium, and nickel, has been reported to be associated with increased lung cancer incidence in those workers. Second, chloromethyl ether or mustard gas, formaldehyde, silica are all known human lung carcinogens. Third, the inhalation of benzo(a)pyrene and other PAH and coke oven emissions has also been linked to the development of lung cancer. Last, radioactive materials have long been recognized as the inducers of lung cancer. Uranium miners are at higher risk of lung cancers, as in the case of victims of atomic bomb explosions at Hiroshima and Nagasaki in Japan. Due to large populations with the possibility for long-term exposure, radon gas inhalation has become a significant concern. Smokers who inhale radon or asbestos fibers may increase their risk of developing lung cancer several folds, suggesting a synergistic interaction between the carcinogens.

In humans, about 90% of lung cancer originates from the cells at the bronchial region (bronchogenic carcinoma). However, a significant increase in peripheral adenocarcinoma has occurred recently. Many studies have been focusing on the potential mechanisms of lung carcinogenesis. So far, DNA damage is thought to be a key mechanism. DNA is believed to interact with an activated carcinogen or its metabolic product, such as alkyldiazonium ions derived from N-nitrosamines. The adduct resulting from the combination seems to play an important role in carcinogenesis. Another potential mechanism of DNA damage is active oxygen species. For example, ionizing radiation leads to the formation of superoxide, which is converted through the action of superoxide dismutase to hydrogen peroxide. Hydroxyl radicals may be formed and may cause DNA strand breaks in the presence of Fe and other transition metals.

Other Lung Diseases

Constrictive bronchiolitis and bronchiolitis obliterans–organizing pneumonia (BOOP) are severe respiratory illnesses producing an obstruction of the small airways. Medical and environmental surveys have revealed the association between toxic inhalations and the development of both BO (e.g., NO_x, chlorine, ammonia, chlorine) and BOOP (e.g., viruses, medicines). Recently, diacetyl, a diketone flavoring agent that is commonly employed for buttery taste, and possibly other flavoring agents, has been implicated as a possible cause of BO. NIOSH investigated several microwave popcorn production facilities between the year 2000 and 2006, and suggested that butter flavoring may be associated with a variety of respiratory diseases, most notably BO. A case report with eight workers from a microwave popcorn plant associated the development of BO with exposure to particulates and a range of organic vapors from flavoring agents where predominant compounds were acetic acid, acetaldhyde, and several ketones (diacetyl, methyl ethyl ketone, acetoin, 2-nonanone) from their monitoring data. The development of BO may be associated with respiratory tract damage induced by diacetyl since studies on rodents exposed to butter flavorings containing diacetyl showed necrosis of the nasal and epithelial injury in pulmonary airways.

10.5 EVALUATION OF TOXIC DAMAGE IN THE RESPIRATORY SYSTEM

Studies in Humans

Since the lung is susceptible to multiple toxicant injuries, tests of pulmonary function evaluation are necessary. As endorsed by the American Thoracic Society, lung function tests should facilitate the description of respiratory dysfunction, assessment of disease severity, and estimation of prognosis in lung disease. Spirometry is the most fundamental element of lung function testing. It directly measures the volume of air exhaled or inhaled by a subject as a function of time. Thus, the reported values from spirometric tests may be measures of volume or flow. Table 10.2 lists the functional values from direct measurement or indirect calculation from values detected with spirometry. The most important ones are the forced vital capacity (FVC), the forced expiratory volume in the first second (FEV1), and the FEV1/FVC ratio.

There are also other tests to evaluate the distribution of ventilation, lung and chest wall compliance, diffusion capacity, and the O_2 and CO_2 contents in the blood. The arterial partial pressures of O_2 and CO_2 are important indices of gas exchange (gases diffuse across the air–blood barrier in the lung). Abnormal partial pressure of gases may be shown when gas exchange is hindered in some situations with the accumulation of fluids or cellular elements in the alveoli (edema, pneumonic infiltrates), thickening of the alveolar wall (fibrosis), insufficient ventilation of alveolar region (emphysema), or insufficient presence of O_2 transport elements (reduced alveolar blood volume or reduced amount of hemoglobin in the blood).

In the clinic, several additional techniques are used to evaluate human pulmonary function. For example, computed tomography (CT) provides detailed roentgenographic information of airways and lung parenchyma. Fiberoptic bronchoscopy has become one of the most valuable tools for detection of lung injury. During the bronchoscopy, saline solutions can be introduced into the lung and then retrieved so that cellular and molecular constituents (broncholveolar

TABLE 10.2 Definition and Interpretation of Common Pulmonary Function Values

Reported Value	Description	Interpretation
Tidal volume (TV)	The volume of air inspired and expired with each breath	It represents the normal volume of the air without extra effort applied
Inspiratory reserve volume (IRV)	The maximum volume that may be inhaled from the resting end-tidal position	It is correlated with function of inspiratory muscles
Expiratory reserve volume (ERV)	The maximum volume that may be exhaled from the resting end-tidal position	It is correlated with function of expiratory muscles
Residual volume (RV)	The volume that remains in the lungs after maximal expiration	An increase in the RV indicates air trapping
Vital capacity (VC)	The maximum amount of air expired after a maximum inspiration. VC = ERV + TV + IRV	Typically preserved in obstruction, but reduced in restriction
Forced vital capacity (FVC)	The maximum amount forcefully exhaled with maximum speed and effort, from maximal inhalation to maximal exhalation	Pattern similar to VC, although more likely to be reduced in obstruction than VC. Used to grade severity of restriction
Forced expiratory volume in one second (FEV1)	Forced expiratory volume in the first second of forced expiratory vital capacity	Reduction typical of medium to large airways obstruction. Used to grade severity of obstruction
FEV1/FVC	Ratio of FEV1 to FVC	Reductions indicative of airway obstruction
Total lung capacity (TLC)	The amount of air in the chest after a maximum inspiration. TLC = RV + ERV + TV + IRV	The lower the TLC, the worse the restriction; The higher the TLC, the worse the obstruction
Functional residual capacity (FRC)	The amount of air remaining in the lungs at the end of tidal position. FRC = RV + ERV	An increase in the FRC indicates hyperinflation
Inspiratory capacity (IC)	The maximum amount of air inspired from the end-expiratory position. IC = TV + IRV	It is correlated with function of inspiratory muscles
Maximum voluntary ventilation (MVV)	Estimate of 1-min maximal air displacement extrapolated from repeated inspiratory and expiratory efforts	Disproportionate reductions relative to FEV1 may indicate upper airways obstruction, muscle weakness, or poor performance
Peak expiratory flow	Maximal sustained airflow achieved during the forcefully exhalation maneuver	Worsening may correlate with asthma exacerbations. Sometimes helpful in assessing subject effort

lavage) are subsequently analyzed. Also, excision of small tissue samples (biopsies) during the bronchoscopy process is an additional diagnostic tool, which is helpful in the evaluation and staging of precancerous and cancerous lesions.

Studies in Animals

Animal Exposure

The inhaled pollutants have been and continue to be extensively studied in experimental animals, such as in rat, mice, guinea pigs, and rabbits. The toxicity studies of airborne pollutants require that special attention be given on generating and delivering well-characterized exposure of test materials to the breathing zone of test subjects. The principal elements of an exposure system are the generation of test atmosphere, the exposure chamber, the monitoring and characterization of the exposure, and the exhaust clean-up. The whole body exposure system works well in its way that the animals are kept within a chamber that is ventilated with a defined test atmosphere. The processes of metering, denseness, and dilution for appropriate concentrations of the exposure are key components. Final concentrations within the chamber are monitored via suitable equipments that appropriate and accurate calibration is essential. In some studies, animals may be exposed to complex mixtures (tobacco smoke, diesel, and gasoline exhaust or residual oil fly ash) instead of particle matters from the ambient air. Compared to nose-only exposure system, which may induce stress to the animals due to confinement during the exposure, whole-body exposure seems less "manipulative" to the mice. The requirement for whole-body exposure chambers include rapid attainment of the desired concentration of toxicants, maintenance of desired toxicant levels homogeneously throughout the chamber, adequate capacity for experimental animals, and minimal accumulation of undesired products associated with animals occupancy.

Pulmonary Function Test

Because experimental animals lack the coordination to maximally inhale or exhale for some pulmonary function tests as in human subjects, conducting pulmonary function tests in

animals poses distinct challenges in some situation. FEV can be obtained in experimental animals only under anesthesia, during which expiration is forced by applying external pressure to the thorax or negative pressure to the airways. However, it is much easier to perform lung compliance test in animals. Compliance is calculated as the slope of the volume–pressure curve, indicating the intrinsic elastic properties of the lung parenchyma and the thoracic cage. All that is needed for the compliance test is cannulation of excised trachea, and attachment to a syringe and manometer to quantify the volume and pressure. The analysis of breathing patterns (respiratory frequency) is another test widely used in animal studies. The main problem is to differentiate sensory upper airway irritants (produced by highly water-soluble irritants such as ammonia, chlorine, and formaldehyde) and pulmonary irritants (produced by less soluble gases such as NO_2 and O_3). Whereas the sensory irritants in the upper airway usually result in slow, deep respiration, pulmonary irritants tend to increase respiratory frequency and decrease minute volume, showing rapid, shallow breathing.

Lung Lavage Assay

Pulmonary edema and/or pulmonary inflammation appear to be obligate early events in acute and chronic lung injuries. The analysis of lavage fluid seems a useful tool to detect respiratory tract toxicity. The lungs of animals are washed with multiple small volumes of isotonic saline and the fluid in saline is collected. The influx of neutrophils, macrophages, or other leukocytes such as lymphocytes or eosinophils into the lavage fluid is the most sensitive and useful sign of inflammatory response in terms of cellular biomarkers. Many secretory cytokines are also measurable, including TNF-α, fibronectin, IL-1, IL-8, monocyte chemoattractant protein (MCP) 1.

Lung Perfusion

The lung perfusion method is applicable to lungs in many laboratory animal species, such as mouse, rat, guinea pig, and rabbit. The lung *in situ* is perfused with blood or a blood substitute through the pulmonary arterial bed. Meanwhile, the lung is actively (through rhythmic inflation–deflation cycles with positive pressure) or passively (by creating negative pressure with an artificial thorax in which the lung is suspended) ventilated. With or without drug treatment, toxic agents can be introduced into the perfusate or the inspired air. The rate of metabolism of drugs, the metabolic activity of the lung, and the effect of drugs can be studied with this technique.

Fixation Technique for Morphological Study

The pathology of acute or chronic injury may be conducted after the examination of the respiratory tract under the microscope. Morphological evaluation is used not only for peripheral lung but also for nasal passages, the larynx, and major airways. For this technique, careful consideration must be given to tissue fixation and preparation. Whereas nasal passages should be flushed with fixative, the lung fixation is done by vascular perfusion with fixatives through the pulmonary artery or by instillation of fixatives through the trachea. Perfusion fixation keeps lining fluid, deposited particles, or cells in the lumen of the airways or the alveoli from their original position. Although trachea instillation may dislodge toxicants, it will keep the alveoli open. The choice of fixatives depends on how the lung will be further analyzed. Formalin-based fixatives are available for routine histopathology, whereas careful selection of the fixatives is required for electron microscopy, immunohistochemistry, and *in situ* hybridization.

In vitro Approaches for Both Human and Animal Study

To study the mechanisms of respiratory injury, *in vitro* studies with samples originally from human or animal tissues are particularly interesting at the tissue and/or cell levels.

Morphologic Techniques

Sections of respiratory tract tissue from humans or animals are widely used in morphological studies. For example, gross pathological changes such as inflammation and presence of cancerous tissue can be detected easily with ordinary sections. Plastic or Epon sections about 1 μm thick are good for proper identification of different cell types lining the airways or alveoli, and for the recognition of cytoplasmic changes in damaged Clara cells. Transmission electron microscopy is essential for ultramicroscopic structural alterations, such as degenerative changes or necrosis of type I epithelial cells or capillary endothelial cells. Confocal microscopy is an ideal tool for three-dimensional reconstruction of normal and damaged lungs, allowing the examination of thick sections and discovery of specific cell types labeled with fluorescent markers. There are additional tools for the investigation of pulmonary injury. With antibodies to a variety of enzymes, mediators, or proteins, immunohistochemistry is especially useful to identify cell types that carry certain enzymes and their anatomical locations. *In situ* hybridization allows one to visualize anatomical sites where a specific gene product is expressed. When damaged by a toxic insult, the respiratory tract is followed by specific cell damage, proliferation, differentiation. Flow cytometry is valuable in the study of these alterations of cell populations prepared from the lung tissue.

Microdissection

Some inhalants act in circumscribed regions of the respiratory system, such as the terminal bronchioles, a region especially rich in the highly metabolically competent Clara cells. After

stripping of the surrounding parenchyma away from small bronchi or terminal bronchiole, the microdissected airways can be used in many studies, including site-specific gene expression and morphological alterations in response to toxicants or drug intervention.

Cell Culture

Many specific pulmonary cell types have been isolated and maintained as primary cell lineages *in vitro*. Alveolar macrophages are commonly used in toxicant exposure experiments with or without drug intervention primarily because they are easy to isolate from human or animal lung lavage fluid or tissue and are closely associated with the development of many diseases. Type I and type II alveolar epithelial cells, fibroblasts, Clara cells, neuroepithelial cells are all available with enzymatic digestion of the lung. However, some critical cellular components and the normal traits of the cell types within the tissue layers may be lost during the process of cell isolation with enzymatic digestion. Other techniques that can maintain the normal function of cell types are appealing. For example, the issue may be settled by tissue culture systems, which have been developed in epithelial cells culture. The surfaces of epithelial cells are exposed to air or a gas phase containing an airborne toxic agent, or particle matters, while the basal portion is bathed by a tissue culture medium. The polarity and differentiation of the epithelial cells are kept with the epithelial cell maintained at the air–liquid interface, and normal function is maximally preserved similar to what is observed *in vivo*. This organotypic cell culture system may stand for a pivotal trend in respiratory toxicity study.

10.6 SUMMARY

The respiratory system poses a unique target for inhaled toxicants. The uniqueness of the respiratory system relates to its physical structure, and location provides a highly efficient method of protection from commonly encountered potential toxicants. The deposited particles are cleared by different mechanisms at different respiratory sites: air filtration, warm and toxic inhalant transport in extrathoracic compartment, mucociliary escalator move-out in tracheobronchial compartment, and phagocyte cell clearance in alveolar compartment. Toxic inhalants include gas and particles. The pattern of toxicity of those gases determines by the sites of gas deposition in the respiratory tract and the condition of the alveolocapillary membrane. Many toxic particles undergo characteristic deposition in certain regions of the respiratory system according to various physicochemical processes, including respiratory pattern and regional characteristics of the respiratory tract.

Inhaled agents exert toxic effect on respiratory system by several distinct pathophysiological mechanisms and lead to many acute or chronic response to lung injury. Acute response primarily includes inflammation, oxidative stress, immune reactivity, alteration of airway function, and acute pulmonary edema. Chronic response mainly includes emphysema, fibrosis, asthma, chronic bronchitis, tumor, and inhibition of lung development. Since the respiratory system is susceptible to multiple toxic injuries, and many workers may suffer from inhaled toxicants-related disease condition, there is a tremendous potential for inhalation toxicity study. Therefore, respiratory toxicology is necessary and critical to be evaluated based on current method systems and investigated thoroughly.

REFERENCES AND SUGGESTED READING

Albert R, Spiro S, Jett J. Clinical respiratory medicine. In: Vassilakopoulos T, Roussos C, editors. *Pulmonary Function Testing, Physiology*. pp. 147–156. In: Hurst JR, Wedzicha JA, editors. *Chronic Obstructive Pulmonary Disease: Management of Chronic Disease, Airway Diseases*. Amsterdam, The Netherland: Elsevier Health Sciences; 2008. pp. 517–522.

Al-Hegelan M, Tighe RM, Castillo C, Hollingsworth JW. Ambient ozone and pulmonary innate immunity. Immunol Res 2011; 49:173–191.

Das R, Blanc PD. Chlorine gas exposure and the lung: a review. Toxicol Ind Health 1993;9:439–455.

Fenn WO, Rahn H. *Handbook of Physiology*. Washington: American Physiology Society; 1964

Gauderman WJ, Avol E, Gilliland F, Vora H, Thomas D, Berhane K, McConnell R, Kuenzli N, Lurmann F, Rappaport E, Margolis H, Bates D, Peters J. The effect of air pollution on lung development from 10 to 18 years of age. N Engl J Med 2004;351: 1057–1067.

Holgate S, Samet J, Koren H, Maynard R. Air pollution and health. In: Paige RC, Plopper CG, editors. *Acute and Chronic Effects of Ozone in Animal Models, Ozone*. pp. 531–558. In: Schlesinger RB, editor. *Toxicology of Sulfur Oxides, Oxides of Nitrogen and Sulfur*. pp. 585–602. San Diego (CA): Academic Press; 1999.

Hollingsworth JW, Maruoka S, Li Z, Potts EN, Brass DM, Garantziotis S, Fong A, Foester WM, Schwartz DA. Ambient ozone primes pulmonary innate immunity in mice. J Immunol 2007;179:4367–4375.

Klaassen CD. Casarett and Doull's toxicology: the basic science of poisons. In: Lehman-McKeeman LD, editor. *Absorption, Distribution, and Excretion of Toxicants, Disposition of Toxicants*. pp. 131–159. In: Witschi HR, Pinkerton, KE, Van Winkle LS, Last JA, editors. *Toxic Responses of the Respiratory System, Target Organ Toxicity*. pp. 609–630. In: Costa DL, editor. *Air Pollution, Environmental Toxicology*. pp. 1119–1156. New York: The McGraw-Hill Companies, Inc; 2007.

Kuempel ED, Wheeler MW, Smith RJ, Vallyathan V, Green FH. Contributions of dust exposure and cigarette smoking to

emphysema severity in coal miners in the United States. Am J Respir Crit Care Med 2009;180:257–264.

Kulkarni N, Pierse N, Rushton L, Grigg J. Carbon in airway macrophages and lung function in children. *N Engl J Med* 2006; 355:21–30.

Marquardt H, Schafer SG, McClellan RO, Welsch F. Toxicology. In: Muhle H, Mcclellan RO, editors. *Respiratory Tract*. pp. 331–348. In: Eyer P, editor. *Gases*. pp. 805–832. San Diego (CA): Academic Press; 1999.

Mauad T, Rivero DH, de Oliveira RC, Lichtenfels AJ, Guimaraes ET, de Andre PA, Kasahara DI, Bueno HM, Saldiva PH. Chronic exposure to ambient levels of urban particles affects mouse lung development. Am J Respir Crit Care Med 2008;178:721–728.

McCreanor J, Cullinan P, Nieuwenhuijsen MJ, Stewart-Evans J, Malliarou E, Jarup L, Harrington R, Svartengren M, Han IK, Ohman-Strickland P, Chung KF, Zhang J. Respiratory effects of exposure to diesel traffic in persons with asthma. N Engl J Med 2007;357:2348–2358.

Pope CA 3rd, Burnett RT, Thun MJ, Calle EE, Krewski D, Ito K, Thruston GD. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. JAMA 2002;287:1132–1141.

Takano H, Yanagisawa R, Ichinose T, Sadakane K, Yoshino S, Yoshikawa T, Morita M. Diesel exhaust particles enhance lung injury related to bacterial endotoxin through expression of proinflammatory cytokines, chemokines, and intercellular adhesion molecule-1. Am J Respir Crit Care Med 2002;165:1329–1335.

Vallyathan V, Castranova V, Shi X. Oxygen/nitrogen radicals: lung injury and disease. In: Rahman I, editor. *Smoking-Induced Inflammation, Injury, and Disease: Molecular Mechanisms*. pp. 211–256. In: Castranova V, Vallyathan V, editors. *Oxygen/ Nitrogen Radicals and Silica-Induced Diseases*. pp. 161–178. New York: Marcel Dekker, Inc; 2004.

William N. Rom. Environmental and occupational medicine. In: Lippmann M, Chen LC, editors. *Particle Deposition and Pulmonary Defense Mechanisms, Organ System: Lung.* pp. 168–186. In: Garay SM, editor. *Pulmonary Function Testing, Organ System: Lung*. pp. 200–236. In: Jalloul AS, Banks DE, editors. *The Health Effects of Silica Exposure, Organ System: Lung*. pp. 365–387. In: Guidotti TL, editor. *Respiratory Tract Irritants, Organ System: Lung*. pp. 543–555. In: Lippmann M, editor. *Ozone, Personal and General Environment*. pp. 1445–1465. In: Schlesinger RB, editor. *Nitrogen Oxides, Personal and General Environment*. pp. 1466–1479. In: Frampton MW, Utell MJ, editors. *Sulfur Dioxide, Personal and General Environment*. pp. 1480–1486. In: Hecht SS, Samet JM, editors. *Cigarette Smoking, Personal and General Environment*. pp. 1521–1551. Philadelphia, PA: Lippincott Williams & Wilkins; 2007.

Williams PL, James RC, Roberts SM. The principles of toxicology: environmental and industrial applications. In: Dallas CE, editor. *Pulmonotoxicity: Toxic Effects in the Lung*. pp. 169–187. New York: John Wiley & Sons, Inc; 2000.

Yadav AK, Bracher A, Doran SF, Leustik M, Squadrito GL, Postlethwait EM, Matalon S. Mechanisms and modification of chlorine-induced lung injury in animals. Proc Am Thorac Soc 2010;7:278–283.

11

IMMUNOTOXICITY: TOXIC EFFECTS ON THE IMMUNE SYSTEM

ERIC S. SOBEL AND STEPHEN M. ROBERTS

Exposure to a variety of chemicals and biological agents has been implicated in the onset of symptoms of immune origin, including acute and chronic respiratory distress, dermal reactions, and manifestations of autoimmune disease. The types of substances associated with immune system effects is extraordinarily diverse and include chemicals found in occupational and environmental settings, infectious materials, certain foods and dietary supplements, and therapeutic agents. As discussed in this chapter, dysregulation of the immune system by toxicants can lead directly to adverse health effects, as well as rendering the body more susceptible to infectious disease and cancer.

The immune system is highly complex, with many facets poorly understood. Because of this, assessment of potential immunotoxic effects of drugs, chemicals, and other agents is not a simple task. Often, measurement of a variety of components of the immune system and/or their functionality is required to gain an appreciation of the likelihood of immune dysfunction from drug or chemical exposure. Increasingly, there is realization that the immune system may be among the most sensitive target organs for toxicity for many chemicals and, as a result, merits special attention.

This chapter discusses the following:

- Basic elements and functioning of the immune system
- Types of immune reactions and disorders
- Clinical tests to detect immunotoxicity
- Tests to detect immunotoxicity in animal models
- Specific chemicals that adversely affect the immune system

11.1 BIOLOGY OF THE IMMUNE RESPONSE

The immune system has evolved primarily to defend the body against the invasion of microorganisms, although normal immune function is important in regulating and sustaining the internal environment as well, such as recognition and removal of malignant cells. There are two types of immunity: natural immunity (also termed *innate immunity*) and acquired immunity (also termed *specific immunity*).

Innate immunity encompasses a wide variety of protective mechanisms, including physical (e.g., barrier activity of skin), chemical (e.g., acidic environment in the stomach), and mechanical (e.g., coordinated movement of cilia in the respiratory tract removing particulate matter). It is also mediated by a number of agents, both secreted (e.g., complement, defensins, type I interferons) and cellular (e.g., neutrophils, macrophages, and natural killer (NK) cells). While originally thought that cells of the innate immune system were activated with little specificity, recent research has identified a number of pattern recognition receptors for ligands that represent potentially dangerous events. These include the recognition of common motifs from exogenous threats, collectively called *PAMPs* (*pathogen-associated molecular patterns*) and endogenous triggers, known as *DAMPs* (*damage-associated molecular patterns*). A family of 11 *TLR* ("*toll-like*" receptors), for example, recognizes specific motifs either not usually seen in mammalian cells or found in unexpected places. These motifs include bacterial lipoproteins, single-stranded and double-stranded RNA, flagellin, and bacterial DNA. *Natural immunity* differs from adaptive immunity in that the receptors are DNA-encoded at the germ line and is rarely enhanced by prior exposure to these substances.

Principles of Toxicology: Environmental and Industrial Applications, Third Edition. Edited by Stephen M. Roberts, Robert C. James and Phillip L. Williams.

Acquired immunity, in contrast, increases in magnitude and specificity with successive exposure to foreign substances. Substances that trigger these specific immune responses are termed *immunogens*, and may be either foreign or endogenous. In many cases, immunogens are proteins or contain a protein as part of the complex. There are two major types of acquired immune responses: humoral immunity and cell-mediated immunity. *Humoral immunity* involves the production of proteins capable of binding to foreign substances. These belong to a special class of proteins called *immunoglobulins* (Igs), and the secreted proteins themselves are called *antibodies*. The substances to which the antibodies bind are called *antigens*. Antibody binding can neutralize toxins, cause agglutination of bacteria and other microorganisms, and lead to precipitation of soluble foreign proteins. By coating the pathogen with antibody that can activate complement, it can also direct cells of the innate immune system to eliminate the pathogen through enhanced phagocytosis—a process known as *opsonization*. Each of these is important in defense of the host. In *cell-mediated immunity*, specialized cells rather than antibodies are responsible for the destruction of foreign cells.

A critical function of the immune system is to effectively distinguish between macromolecules that belong, or do not belong, in the body. The specific immune response is believed to be highly individualistic, a process which defines "self" while also defending the organism against "nonself." This is evident by the response to certain environmental toxicants, to allergens or antigens, and the specific rejection of allografts. Recognition of "self" is known to be guided, in part, by genetic variations in proteins of the class I and II *major histocompatibility complex* (MHC). Initially, the ability of the immune system to differentiate "self" from "nonself" is the result of a rigorous selection process. During maturation, the system must ignore an infinite variety of self-molecules and yet be primed and ready to respond to an array of exogenous antigens. Immunomodulatory control mechanisms lead to immune tolerance of self and carefully orchestrate the immune response to targets and removal of foreign macromolecules and cells. These control mechanisms arise from interactions among the several different cell types with roles in proper immune function.

Lymphocytes are considered to be the major cells involved in a specific immune response in humans. They are derived from pluripotent stem cells and undergo an orderly differentiation and maturation process to become T cells or B cells, with critical functional roles in the host defense. T-cell development occurs primarily in the thymus, where certain cell surface protein markers are acquired during the selection and differentiation process. Collectively, protein markers expressed on cells of the immune system are called CD antigens (for *cluster of differentiation*), and well over 300 different CD antigens have been identified in humans. The presence of combinations of CD antigens, detectable by immunofluorescence, has been used to positively identify immunocytes. In general, mature T cells exiting the thymus and present in the circulation and secondary lymphoid tissue (lymph nodes and spleen) are characterized by the presence of $CD3^+$ and either $CD4^+$ or $CD8^+$ surface markers and are devoid of surface or cytoplasmic Ig. CD3 is a complex that associates with a T cell receptor (TCR), and it is the TCR that has been the target of the selection process, mainly to avoid those receptors that either cannot bind to self MHC (and are therefore useless) or bind too avidly to self MHC (and are therefore likely self-reactive). There are various subtypes of T cells, such as T-helper (T_H) cells, regulatory T (T_{reg}) cells, and cytotoxic T cells (T_C). T_H and most T_{reg} lymphocytes carry the $CD4^+$ marker, while T_C lymphocytes have the $CD8^+$ marker. Together, these T-lymphocyte populations play a vital role in initiating and regulating the immune response.

Human B cells develop from stem cells in the fetal liver and, after birth, B-cell development occurs principally in the bone marrow. B-cell development and maturation are characterized by class-specific Ig expression on the cell surface. B cells also undergo a selection process to avoid self-reactivity, although it is not as rigorous as for T cells. Monoclonal reagents can identify the Ig expressed on the surface of B cells. Immunophenotypic characterization of cells via these and other CD markers has proved to be invaluable in certain clinical situations (such as identifying the type of lymphoma) and a useful research tool. B cells play an important role in recognition of antigens and are responsible for antibody production.

Another important cell in the specific immune response is the "*professional*" *antigen presenting cell* (APC). These cells make first contact with the antigen and may also process the antigen; that is, modify it in such a way as to enable its recognition by T cells. This category of cells is defined more by function than cell type. There are three types of professional APCs: peripheral blood monocytes (which can differentiate into tissue macrophages), B cells, and dendritic cells. Professional APCs are distinguished by their ability to express both MHC class I and II; virtually all other nucleated cells express only class I. Peptides processed and bound to class II are generally of extracellular origin, while peptides bound to class I are derived from the cytoplasm of that cell, providing a pathway for internal surveillance against important pathogens, such as viruses. Dendritic cells are responsible for initiating most primary (*naïve*) T cell responses, and are unique in their ability to process exogenous antigen by both the class I and II pathways (*cross-presentation*).

In order for the APC to present the antigen to T cells, the antigen must have a polypeptide component to it which is processed, or partially digested by the APC, and sequences from that peptide are "presented" on its cell surface bound to an MHC class I or class II molecule. Presentation of antigen to B cells does not require this processing, and in fact B cells

are capable of recognizing antigens directly. These antigens interact with Igs serving as receptors on the cell surface of B-cell clones. Every Ig molecule expressed by an individual B cell has identical specificity, and these Igs can be quite specific in terms of the antigens with which they will interact. Thus, a particular antigen may interact with only one or a few B cell clones, a critical aspect in creating a specific immune response. When the antigen binds to an Ig receptor on the B cell surface, the antigen–receptor complex migrates to one pole of the cell and is internalized within the cell. While the Ig receptor can bind to virtually any structure, for the B cell to become activated and productively engage specific T cells, the antigenic complex must also have a polypeptide component that can be processed and antigenic peptides displayed via the MHC class II pathway.

Naïve T-cell activation requires at least two signals. The first signal results from an interaction between the T-cell receptor and the MHC/peptide complex and provides the specificity. The second signal is the product of interactions between co-stimulatory receptors constitutively expressed by T cells and their ligands, usually inducibly expressed by the APCs. Expression of the co-stimulatory ligands is induced by the danger signals of the innate immune system. To facilitate the interactions between the rare T cell specific for a peptide derived from the pathogen and the APC presenting those peptides, dendritic cells activated by danger signals at the site of contact take up antigen and migrate (along with free antigen) via the lymphatic system to the nearest draining lymph node. There, they encounter naïve T cells that are programmed to continuously circulate through lymph nodes. The architecture of the lymph node promotes efficient cell-to-cell contact and intercellular signaling.

In general, an effective immune response requires the differentiation and proliferation of specific subsets of T_H cells (T_H1, T_H2, and T_H17 cells) which respond to and, in turn, secrete different cytokines in a complex interplay of positive and negative feedback loops. *Cytokines* are low-molecular-weight proteins that mediate communication between cell populations, and the net product of these influences at the initiation of the primary response leads to activation of a master transcriptional factor. A partial list and functional classification of cytokines is shown in Table 11.1. T_H1 cells (master transcriptional factor Tbet) are involved in the activation of macrophages by INF-γ, secrete tumor necrosis factor (TNF), and mediate delayed-type hypersensitivity responses. The most critical function of T_H2 cells (master transcriptional factor GATA-3) is to stimulate B cells to produce antibody, but they also secrete cytokines (specifically, interleukins, designated IL) that may regulate mast cells (IL3, IL4, and IL10), eosinophils (IL5), and IgE (IL4) responses in allergic diseases. Of the several factors known to participate in immunomodulation, IL4 and IL10 are particularly noted to upregulate the humoral response while suppressing the cell-mediated response (see in the following text for more discussion of humoral vs. cell-mediated immunity). IL13, which is produced by activation of T cells (Table 11.1) and shares many of the properties of IL4, also suppresses cell-mediated immune responses and the production of proinflammatory cytokines (IL1, IL6, IL8, IL10, IL12, and TNF). More recently, T_H17 cells (named for the effector cytokine IL-17 and differentiated by the master transcriptional factor RORγ) have been shown to be important mediators for many chronic inflammatory responses. An effective cytotoxic T-cell response to a novel pathogen typically also requires prior or coincident activation of the APC (typically a DC) by T_H cells.

Similar to naïve T cells, naïve B cells circulate through lymph nodes via the blood. Within the lymph node, the rare B cell with Ig receptor specificity for an *epitope* (antigenic domain) present in the antigen can bind antigen with high affinity and process the protein component via the MHC class II pathway. When an activated T_H cell binds to the antigenic peptide–MHC complex on that B cell, the B cell is stimulated to replicate and differentiate into an antibody secreting *plasma cell*. This B-cell clonal expansion occurs in the germinal center of the lymph node in a highly orchestrated fashion that leads to survival and proliferation of the B-cell clones with the highest affinity for the epitope that originally selected that B cell. This process of *affinity maturation* occurs through somatic hypermutation (programmed mutations in the DNA of the regions of the Ig genes coding for the binding site) and selection. Through this mechanism, the immune system is able to produce the necessary quantities of antibodies targeting specific molecules (antigens) regarded as foreign. The synthesis of the antibody is tightly regulated; however, the proliferation of plasma cells and antibody synthesis are controlled by cytokines and interactions with T cells. Regulatory T cells (T_{reg}; master transcriptional factor FoxP3) function to suppress the immune response. Control of the immune response is achieved by balancing the stimulatory and inhibitory effects of T cells and various cytokines.

After an encounter with an antigen, antigen-specific T cells undergo rapid proliferation and differentiation in secondary lymphoid tissue into "armed effector" and "memory" cells. The specific signals that determine whether an individual T cell becomes an effector or memory cell is not yet well-understood. However, with clearance of the antigen, there is an active down-regulatory mechanism that rids the immune system of now unneeded armed effector T cells but retains an expanded population of antigen-specific memory T cells. It is these memory T cells that are responsible for the stronger and more rapid response to subsequent encounters with the antigen that characterizes adaptive immunity. Moreover, subsequent responses no longer absolutely require a co-stimulatory signal. An analogous process occurs for B cells, with the outcome being either

TABLE 11.1 Major Cytokines and their Functions

Cytokine	Produced By	Function(s)
IL1 (IL1-α and IL1-β)	Macrophages, epithelial cells	Variety of effects, including neutrophil and macrophage activation, T- and B-cell chemotaxis, and increased IL2 and IL6 production
IL2	T cells	Stimulates replication of T cells, NK cells, and B cells
IL3	T cells, thymic epithelial cells	Involved in regulation of progenitor cells for several different cell types, including granulocytes, macrophages, T cells, and B cells
IL4	Activated T cells, mast cells	Differentiates T_H2 cells; activates B cells and induces class switching to IgE
IL5	T cells and activated B cells	Increases secretion of immune globulins by B cells; eosinophil growth
IL6	T and B cells, macrophages	Important in inflammatory reactions and in differentiation of B cells into Ig-secreting cells; fever
IL7	Bone-marrow stromal cells	Growth of pre-B cells and pre-T cells
IL8	Activated monocytes and macrophages	Activates neutrophils; important for chemotaxis of neutrophils and lymphocytes
IL10	Monocytes	Inhibits inflammatory and immune responses
IL12	Macrophages, dendritic cells	Activates NK cells; induces differentiation to T_H1 cells
IL17	Activated T cells	Stimulates fibroblasts, epithelial cells, and keratinocytes to secrete cytokines that recruit inflammatory cells
IL23	Dendritic cells	In combination with other cytokines, stimulates differentiation of naïve T cells to become T_H17 cells
TNF-α	Variety of cells, primarily activated macrophages	Important in inflammatory responses; effects similar to IL1
TGF-β	Variety of cells	Inhibits T-cell proliferation and suppresses inflammatory responses
TNF-β	Activated CD4+cells(T_H)	Important in mediating cytotoxic immune responses, cell lysis
Interferon-α Interferon-β	Leukocytes and dendritic cells (IFN-α), fibroblasts (IFN-β)	Increased MHC class I expression; inhibits viral protein synthesis
Interferon-γ	T cells, natural killer cells	Macrophage activation; suppresses T_H2

differentiation into antibody-secreting plasma cells or memory B cells. Like memory T cells, memory B cells are responsible for the brisker and higher-titer antibody response seen on subsequent antigen encounter. Also similar to armed effector T cells, the majority of the plasma cells survive only for a few weeks. However, some migrate to the bone marrow and successfully compete for limited niches. There, they receive survival signals and become long-lived plasma cells. Because immune responses to viruses or immunization encountered in childhood generally result in lifelong immunity, it has been assumed that "memory cells" last indefinitely following a single antigen contact. However, evidence suggests that the life span of memory cells may be related to repeat contact with antigen.

In order to be recognized by the immune system, antigens must be of appreciable size. Some of the smallest antigens, for example, are natural substances with molecular weights in the low thousands. There are circumstances where much smaller molecules can elicit an immune response, but this requires the participation of a large molecule to serve as a carrier. For example, some metals, drugs, and organic environmental and occupational chemicals too small to be recognized by the immune system can become antigenic when bound to a macromolecule, particularly proteins. Once the immune response has been initiated, antibodies will recognize and bind the small molecule even when it is not bound to the carrier molecule. In situations such as this, the small molecule is called a *hapten*.

The antibodies themselves are glycoproteins, the basic unit of which consists of two pairs of peptide chains (see Figure 11.1) connected by disulfide bonds. The longer peptide chain is termed the *heavy* (or H) chain and the shorter is the *light* (or L) chain. There are five main types of antibodies, or immunoglobulins (Igs): IgG, IgM, IgA, IgE, and IgD. They differ both in structure and function. IgG is present in the greatest concentration in serum, has a molecular weight of around 150,000 (there are four subtypes of somewhat different sizes), and is important in secondary immune responses. They distribute into the extracellular space throughout most of the body. IgM is a primary response antibody, meaning that it is increased very early in an immune response, especially on the naïve encounter. IgM is much

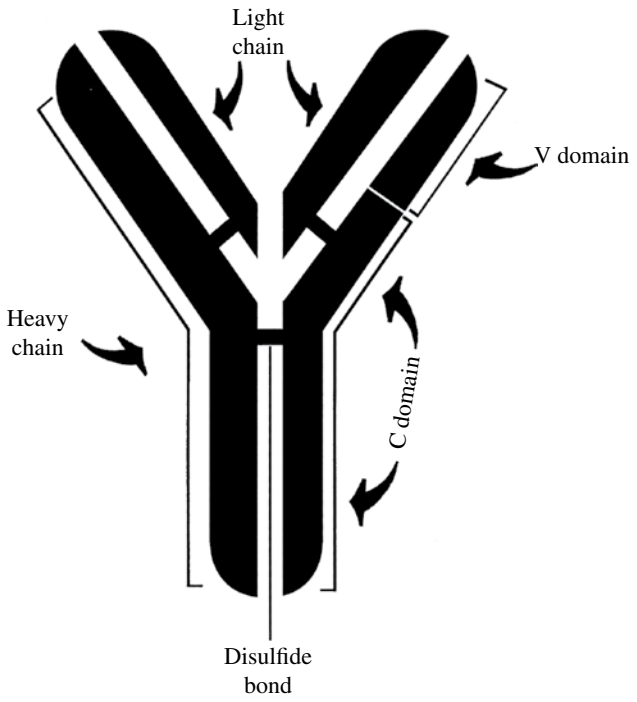


FIGURE 11.1 Light and heavy chain structure of IgG. IgG illustrates the basic structure of antibody proteins, which consists of two long, heavy chains and two shorter, light chains held together by disulfide bonds. Composition of C domains is relatively constant, while V domain varies, creating the binding specificity characteristic of antibodies.

larger than the other Igs, consisting of five sets of heavy/light-chain pairs bound together at a single point with another peptide (the J chain). Its molecular weight is about 970,000. Its size largely confines it to the vascular space, and its multivalency allows it to bind to target even when affinity is low. It is these properties that make it useful as a first responder of the adaptive immune system in eliminating blood-borne pathogens. IgA may exist as a monomer (one basic unit of two pairs of H and L chains) or as a dimer (two basic units bound together with a J chain). The monomeric IgA is the primary Ig found in blood, although present at low concentrations. The dimeric form, with a molecular weight of 385,000, is actively transported to mucosal surfaces, where it serves to prevent binding of certain pathogens to mucosal surfaces. Its dimeric structure resists degradation in that challenging environment. IgD has a molecular weight of about 184,000, and is present in very low concentrations in serum. Its function is unclear, but it may play a role in B-cell differentiation. IgE is slightly larger than IgG (molecular weight of 188,000) and is normally present in vanishingly low concentrations in serum. It binds through high-affinity receptors on mast cells strategically located skin and mucosal barriers and arms them to respond to parasites. It is the primary antibody involved in immediate hypersensitivity reactions.

In cell-mediated immunity, cells carrying the antigen on their surface are attacked directly by cytotoxic T cells (T_C) or other cell types such as NK cells. In the case of T_C cells, recognition of cells to be destroyed is through interaction between cytoplasmically derived processed antigen in conjunction with the nearly ubiquitously expressed MHC class I molecules on the target cell surface and an antigen receptor on the T_C. Activated cytotoxic T cells can kill appropriate targets without additional signaling. NK cells have a set of receptors that can target cells that fail to express class I, reducing the ability of viruses to evade the immune system by preventing class I expression.

The importance of the interplay and integration between the innate and acquired immune system is hard to overemphasize. Signaling through PAMPs results in maturation of dendritic cells and converting them into APCs that can initiate a naïve T-cell response. Conversely, antibodies and Th1 and Th17 T cells greatly enhance the ability of cells of the innate immune system to target pathogens. Finally, although autoimmune diseases are caused by a failure of the adaptive immune system to maintain tolerance to self (see Section "Autoimmunity"), the majority of susceptibility genes thus far identified are of the innate immune system.

11.2 TYPES OF IMMUNE REACTIONS AND DISORDERS

Interactions of toxicants with the immune system may result in undesirable effects of three principal types: those manifested as (i) a hypersensitivity (allergic) reaction, (ii) immunosuppression, or (iii) autoimmunity. Each is discussed as follows.

Allergic Reactions

Allergic reactions are divided into four classes:

Type I. Type I immune response is limited to IgE-mediated hypersensitivity reaction. This reaction involves an initial exposure in which clinical symptoms are generally absent (sensitization), followed by reexposure that can elicit a strong allergic reaction. In type I immune responses, antigen interacts with IgE antibodies passively bound to mast cells. On binding of antigen to the IgE, the mast cells release histamine and serotonin, which are responsible for many of the immediate symptoms of an allergic reaction such as upper respiratory tract congestion and hives. In a severe reaction, termed *anaphylaxis*, histamine and serotonin release can cause vasodilation leading to vasomotor collapse, and bronchiolar constriction making breathing difficult. This type of reaction has occurred following the administration of a number of different

drugs and diagnostic agents, hormones, and a variety of sulfating agents (e.g., sodium bisulfite and sodium metabisulfite).

Type II. Type II reaction is believed to be the result of the binding of a drug or chemical to a cell surface, followed by a specific antibody-mediated cytotoxicity that is directed at the agent (drug or chemical) or at the cell membrane that has been altered by the compound. Clearance is typically mediated by NK cells or phagocytic cells of the innate immune system. Under some circumstances, immune complexes may become absorbed to a cell surface (erythrocytes, thrombocytes, or granulocytes) resulting in a complement-mediated cytotoxic response, leading to induction of immune hemolytic anemia, thrombocytopenia or granulocytopenia. It is also possible for antibodies binding to a cell surface molecule to act as an agonist rather than directing the targeted cell to destruction. Chronic urticaria (recurrent hives) is an example of this mechanism.

Type III. Soluble immune complexes consisting of a drug or chemical hapten (plus carrier molecule) and its specific antibody plus complement components are primarily responsible for immune complex disease. A particular form of immune complex disease arising from injection of an antigen is called *serum sickness syndrome*. Clinically, a type III reaction may be characterized by the onset of fever and the occurrence of a rash that may include purpura (bruise-like lesion caused by bleeding into tissue adjacent to damaged capillaries) and/or urticaria (hives). The immunopathology includes the deposition of immune complexes in areas such as blood vessel walls, joints, and renal glomeruli, typically with the activation of complement. Some of the signs and symptoms associated with drug-related lupus (a systemic autoimmune disease) may be included under type III reactions.

Type IV. These reactions involve cell-mediated and/or delayed-type hypersensitivity responses. The expression of type IV reactions requires prior exposure to the agent and T-cell sensitization. A special subpopulation of effector memory T cells (T_{EM}) appears to be responsible for this reaction. The T_{EM} cells circulate in tissues and respond to processed peptide antigens bound to MHC class II, releasing cytokines and attracting macrophages to the site. The reaction is termed delayed because the inflammatory reaction may not peak for 24–48 h, as opposed to responses occurring within a few minutes to a few hours with other reaction types. These reactions are usually seen after the use of certain drugs or exposure to some chemicals. In a variant of this, very lipid-soluble allergens (e.g., pentadecacatechol present in poison ivy) penetrate into cells and bind and modify amino acid residues on cytosolic proteins. These modified proteins can be alternatively processed by the class I antigen processing pathway, leading to novel peptides that appear foreign (altered self) and stimulating a cytotoxic T-cell response. Finally, the type I immediate hypersensitivity response is ultimately dependent upon a T_H2 response that led to development of IgE antibodies bound to the mast cells.

Immunosuppression and Immune Enhancement

Impairment of one or more components of the immune system from drug to chemical exposure can lead to loss of immune function, or *immunosuppression*. Clinically, this is manifested primarily as increased susceptibility to infectious disease, although diminished immune function has been associated with an increased vulnerability to cancer, presumably by impairing immune surveillance and removal of malignant cells. In certain situations immunosuppression is intentionally induced via drug therapy to prevent rejection of transplants. Agents employed for this purpose are diverse, and several potential mechanisms are involved, including inhibition of cytokine production (e.g., corticosteroids, cyclosporine) and lymphocyte proliferation (e.g., azathioprine). Most of the evidence that environmental and occupational chemicals suppress immune responses is derived from animal studies, and while the same principles likely apply to humans as well, there are few clear examples in the clinical literature of immunosuppression from chemical exposure other than that from intentional treatment with immunosuppressive drugs.

The opposite reaction, immunological enhancement, is also possible, and several natural and synthetic agents have been shown to increase immune responsiveness under experimental conditions. Examples of agents that increase immune reactivity include the bacillus Calmette-Guérin (BCG), alum (aluminum potassium sulfate or aluminum hydroxide), bacterial lipopolysaccharides and peptidoglycans, a variety of synthetic polymers, and the antiparasitic drug Levamisole (phenylimidazolethioazole). The imidazoquinoline imiquimod enhances immune responses by stimulating the TLR7 and TLR8 receptors for ssRNA (part of the innate immune response to viral infections) and has been used clinically in the treatment of cutaneous neoplasms.

Autoimmunity

Autoimmunity is defined as an adaptive immune response directed to self-tissue. While the response may be cellular and/or humoral, autoimmunity is most clearly evident when autoantibodies, including those to nuclear macromolecules, can be identified. Studies of drug-related autoimmunity in humans have provided some of the best examples of this type of reaction. Although there are many types of autoimmune disease, the most common autoimmune syndrome produced by drugs is one resembling systemic lupus erythematosus

(SLE). Clinical signs and symptoms of so-called *drug-induced lupus* are not identical to idiopathic SLE, however. Both can be characterized by arthralgia (joint pain), skin rashes, inflammation of serosal surfaces (lining of lungs and heart), and the appearance of antinuclear antibodies in the blood. However, the pattern of antinuclear antibodies is somewhat different and renal and CNS complications are common in idiopathic SLE but typically absent in drug-induced lupus. Symptoms of drug-induced lupus generally subside after the drug is withdrawn. While intraperitoneal exposure to the hydrocarbon pristane (tetramethylpentadecane or TMPD) results in the development of a lupus-like syndrome in every strain of mouse so far tested, demonstration of autoimmune responses from environmental exposure to chemicals (other than drugs) in humans has been difficult. In part, this is because of problems identifying etiologic agents in retrospective studies of patients developing autoimmune disease, or even clearly documenting prior exposure, which may be remote. One concern is that some chemicals, rather than causing de novo autoimmunity, may exacerbate underlying autoimmune disease (e.g., SLE), rendering symptomatic a patient with subclinical disease or increasing the duration or severity of symptoms in those with active disease. Unfortunately, differentiating the effects of chemical exposure from progression of the underlying disease is difficult or impossible in practice. Understanding of autoimmune consequences of chemical exposure is further hampered by the general lack of satisfactory animal models for many human autoimmune diseases—the results obtained in laboratory animals seldom correspond exactly to observations in humans.

11.3 CLINICAL TESTS FOR DETECTING IMMUNOTOXICITY

In the clinical setting, the use and proper interpretation of immunologic laboratory tests can be important in establishing a differential diagnosis in a patient who has been exposed to an immunotoxic agent. Immune system testing for diagnostic purposes can be challenging, however, because of the complexity of the immune system and difficulty in establishing normal values for many of the tests. When immune dysfunction from chemical exposure is suspected, it is important to be sure that the patient is free from infectious disease and not taking medications that can influence immune function—obvious confounders to interpretation of any immune tests. Also, it is important to recognize that many immune parameters, such as lymphocyte subpopulation counts, can vary normally by age and gender, making the use of appropriate controls essential for proper interpretation of results. Finally, temporal variations in most tests are common. In order to demonstrate that an abnormality exists, it is usually advisable to repeat the test one or more times to insure that a consistent result is obtained.

Some of the laboratory tests available provide information relevant to assessing humoral immunity, others are useful in evaluating cellular immunity, and some can provide insight regarding both. Because the immune system is very complex and a virtually limitless number of different tests could be devised, the strategy has typically been to evaluate clinically relevant parameters that represent the results of an integrated immunological response and then to try to dissect that response if an effect is seen. Described in the following are examples of assays commonly used in the evaluation of individuals exposed to chemicals in the environment or workplace.

Immunoglobulin Concentrations

The concentrations in serum of each Ig can be determined with the exception of IgD, which exists primarily on cell surfaces. Single-radial diffusion is commonly employed for most Igs, although enzyme-linked immunosorbent assay (ELISA) or radioimmunoassay (RIA) is often needed to measure the low concentrations of IgE typically present. Diminished immunoglobulin concentrations, either in total or of specific classes, may suggest immunodeficiency, but is not sufficient to establish a diagnosis. Conversely, Igs within normal limits do not necessarily indicate immunocompetence. There may be defects in subtypes of Igs not quantified by the assay, and patients with normal or high values may nonetheless exhibit increased susceptibility to disease. Ig values may be profoundly influenced by viral or bacterial infections and the presence of some drugs.

T- and B-Cell Concentrations

Immunotyping of T- and B-cell subsets by ethidium bromide and cytofluorometry techniques is used by many laboratories for screening studies of chemical-related injury. Concentrations of B cells, either in absolute terms or as a percentage of peripheral blood lymphocytes, can be expressed, and the distribution of B cells expressing different Ig types (IgM, IgG, IgA) can be measured. Some studies have sought to evaluate a potential immunosuppressive effect through measurement of the ratio of T_H to T_S lymphocytes in peripheral blood, using the $CD4^+$ marker to indicate T_H cells and the $CD8^+$ marker for T_S cells. As discussed earlier, these markers are not specific for T_H and T_S cells, however, and interpretation of a decreased $CD4^+$- to-$CD8^+$ ratio as a loss of T help relative to T suppression is an oversimplification. A significant reduction in $CD4^+$ cells is associated with several immunodeficient states (e.g., in patients with AIDS, undergoing radiotherapy or chemotherapy), implying that diminished $CD4^+$ is indicative of impaired antibody production. This assumption is not infallible, however, because there are also circumstances in which $CD4^+$ cells may be

reduced without loss of antibody production. Significant changes in absolute or relative concentrations of lymphocyte subsets may be suggestive of immunotoxic effects from chemical exposure, but are not, by themselves, reliable indicators of compromised function.

Cutaneous Anergy

Anergy is a generalized clinical condition of nonresponsiveness to ubiquitous skin test antigens that is frequently observed in patients who are immunosuppressed. Cutaneous anergy may suggest functional impairment or abnormalities of the cellular immune system. The most cost-effective method for evaluation of cutaneous anergy is the use of a battery of attenuated, premeasured, and well-standardized ubiquitous antigens that are available from commercial sources. The assessment of a person who is thought to be immunologically suppressed due to exposure to an environmental chemical can be attained within 48 h through the use of these antigens. The intradermal skin test antigens frequently used to measure cellular delayed hypersensitivity are: tetanus toxoid, diphtheria toxoid, *Streptococcus* (group C), old tuberculin (PPD), *Candida albicans*, *Trichophyton mentagrophytes*, and *Proteus mirabilis*. Measurement of specific IgG antibodies to diphtheria and tetanus toxoids in serum at 2 weeks following booster immunization is also useful in assessing the ability for forming antibodies to protein antigens.

In Vitro Tests

Functional capabilities of lymphocytes can be evaluated by taking a blood sample and performing a variety of tests *in vitro*. In general, these tests involve isolating lymphocytes from a blood sample, placing them in culture, and exposing them to a stimulatory agent. The ability of the cells to proliferate in response to the stimulus and, in the case of B cells, to synthesize Igs, can be measured. For example, treatment of peripheral blood lymphocytes with pokeweed mitogen (PWM) normally produces cellular proliferation and increased Ig synthesis. This response requires both T_H and B cells, and provides an indication of the capability of these two cells to interact properly and of B cells to produce Igs. Lipopolysaccharide (LPS) is a mitogen effective selectively on B cells, while phytohemaglutinin (PHA) and concanavalin A (con A) are selective T-cell mitogens. Other stimulants to lymphocyte activation can be used, such as tetanus toxoid, diphtheria toxoid, *Candida*, and PPD, if the subject has been previously exposed to these. The rapid cell division characteristic of a normal response to these mitogens is typically assessed by measuring incorporation of ^{3}H-thymidine into DNA of the cells. Other endpoints of stimulation, such as increased expression of IL2 receptors on T cells, can also be evaluated. The results of these tests are particularly prone to variability, and the tests should be repeated on several occasions in order to demonstrate an abnormal response.

In the mixed-lymphocyte reaction (MLR) test, lymphocytes from the test subject and another individual are mixed. Normally, contact with the allogeneic (i.e., expressing alternate alleles of MHC) lymphocytes will cause the test subject's lymphocytes to become activated and proliferate. To conduct this assay, the target lymphocytes are rendered incapable of replication, often by irradiation or by treatment with mitomycin C. Test subject lymphocytes are then added, and the rate of their replication is evaluated by measuring incorporation of ^{3}H-thymidine. The cytotoxic lymphocyte (CTL) assay takes the lymphocyte interactions one step further to evaluate the ability of cytotoxic T cells (T_C) to destroy target cells. After incubation of the test subject and target lymphocytes, the subject T_C are isolated, washed, and reincubated with target lymphocytes preloaded with radioactive ^{51}Cr. As the target cells are destroyed, ^{51}Cr is released into the medium and can be measured, providing an index of cytotoxic capabilities of the T_C lymphocytes.

Fluorescent Antinuclear Antibody Assay

The indirect immunofluorescence antinuclear antibody assay (FANA) may be the initial screening test used to show autoimmunity. However, several FANA patterns are recognized in various connective-tissue diseases and some low-titer staining patterns have also been reported in sera from persons exposed to environmental agents. The following staining patterns may be observed:

1. The *diffuse (homogenous) staining pattern*, which is usually associated with antibody directed to DNA–histone or histone subfractions. This staining pattern is frequently found in sera from patients receiving chronic treatment with procainamide, hydralazine, isoniazid, anticonvulsant drugs, and some environmental chemical agents.
2. A *peripheral (rim) pattern*, which is attributed to antibody reacting with native DNA and soluble DNA–histone complexes. This staining pattern is frequently seen in sera from patients with SLE (>95%).
3. *Speckled FANA staining*, which is usually attributed to antibodies reacting with saline-soluble antigens. These antibodies are directed to nonhistone antigens and include Sm, ribonucleoprotein, SS-A/Ro, SS-B/La, PM-1, and SCL-70. While these staining patterns frequently occur in patients with mixed connective tissue diseases, including Sjögren's syndrome, polymyositis, and progressive systemic sclerosis, they have also been found in sera from persons exposed to immunotoxic agents.

4. The *nucleolar staining pattern*, which has been restricted to antibodies reactive with nucleolar RNA. This pattern is associated with a particular form of systemic sclerosis (progressive systemic sclerosis).

11.4 TESTS FOR DETECTING IMMUNOTOXICITY IN ANIMAL MODELS

For most chemicals, an assessment of their potential to produce immunotoxicity in humans is based on testing in animals. Many of the tests used in animal studies are the same as, or at least analogous to, those available for clinical assessment described earlier. However, studies in animals offer the opportunity to evaluate directly toxic endpoints difficult or impossible to assess clinically, such as the development of immunopathology or loss of resistance to infectious disease.

Standard toxicity tests typically include measurement of endpoints that can provide an indication whether immunotoxicity is occurring and thus serve as an initial screen. Examples of effects that could be interpreted as suggesting immunotoxicity include changes in hematology (e.g., altered white blood cell counts), lymphoid organ weights, serum globulin concentrations, and incidences of infections or tumors. Careful examination of histopathology of the immune system (e.g., lymph nodes, spleen, thymus) can also contribute to an initial evaluation of potential immunotoxicity.

When immunotoxicity of a drug or chemical is suspected, more definitive insight can be gained by conducting a series of functional tests. Humoral immunity can be assessed with a plaque-forming cell (PFC) assay. In this assay, the test animal is injected with sheep red blood cells (SRBCs) as the source of antigen. Four days later the spleen is removed, and cells isolated from the spleen are cultured with intact SRBCs. B cells producing IgM directed to SRBC antigens result in lysis of the red cells, producing clear areas in the culture called plaques. The number of plaques (per spleen or per million spleen cells) provides an indication of the ability of splenic cells to synthesize and secrete antigen-specific antibodies. This, in turn, offers information regarding the ability of the immune system to mount a primary (IgM-mediated) response. Humoral immunity can also be evaluated through measurement of specific serum antibody titers.

Cell-mediated immunity is commonly evaluated by measuring the responsiveness of peripheral blood T and B lymphocytes to mitogens (such as concanavalin A). The MLR and CTL assays are also used to assess cell-mediated immunity in experimental animals. Nonspecific immunity is evaluated by measuring NK cell function. These tests are essentially identical to the *in vitro* methods described earlier for clinical assessment of potential immunotoxicity in humans.

If warranted, more detailed tests are available to further characterize drug and chemical effects on the immune system. For example, if disturbance in the numbers of immunocytes is suggested by hematology findings, the abundance of individual T- and B-cell types in the spleen or blood can be measured using reagents that detect specific cell surface antigens. In the assessment of humoral immunity, an abnormal primary response (IgM-mediated) to SRBCs detected in the PFC assay might lead to an evaluation of the secondary response (IgG-mediated) to SRBCs. Evidence of altered cell-mediated immunity could lead to expanded tests of T-lymphocyte cytotoxicity, commonly using tumor cells as targets. Additional evaluation could also include an assessment of delayed-type hypersensitivity response. Evaluation of nonspecific immunity may be extended to include enumeration of macrophages and tests of their function. For functional tests, macrophages are typically taken from the peritoneal or alveolar space of test animals, cultured, and examined for phagocytic activity, secretion of cytokines, and/or production of reactive oxygen or nitrogen species. The ability of macrophages in culture to phagocytize foreign materials is typically examined using light microscopy, with either biological (e.g., SRBCs or bacteria) or nonbiological materials (e.g., fluorescent beads) as targets. On activation, macrophages normally release specific cytokines (e.g., TNF-α and IL2), as well as reactive oxygen and nitrogen. Cytokine production by activated macrophages in culture can be measured by ELISA (enzyme-linked immunosorbent assay) using antibodies directed to specific cytokines, or by ELISPOT, which is capable of identifying the numbers of cells producing specific cytokines. Several techniques are available for quantitating reactive oxygen and nitrogen species.

When immunosuppression (or, less commonly, immunostimulation) is suspected, one of the most direct means to test overall immune competence is through a *host resistance model* (also sometimes called a *host susceptibility model*). With this model, the ability of the animal to withstand an immune challenge is assessed with and without exposure to the drug or chemical. Immune challenge can take the form of an infectious microorganism or a syngeneic tumor. A variety of types of infectious microorganisms are used for these tests, including viruses, bacteria, yeast, fungi, and parasites. Syngeneic tumor lines are derived from the same strain and species as the test animal, requiring their recognition as tumor cells, and not simply a source of foreign protein. Examples of microorganisms and tumor cell lines used for host resistance models are provided in Table 11.2. Many of these agents are human pathogens, and this type of test arguably provides the best direct evidence of the ability of a drug or chemical to produce clinically relevant immune suppression or stimulation.

TABLE 11.2 Examples of Agents Used for Immune Challenge in Host Resistance Tests

Type of Agent	Name	Typical Exposure Route
Virus	Cytomegalovirus	Intraperitoneal or intratracheal administration
	Herpes simplex virus type 2	Intraperitoneal, intravenous, or intravaginal administration
	Influenza virus	Intranasal administration
Bacteria	*Corynebacterium parvum*	Injected intravenously
	Listeria monocytogenes	Injected intravenously
	Pseudomonas aeruginosa	Injected intravenously
	Streptococcus pneumoniae	Injected intravenously
Parasites	*Plasmodium* species	Intravenous or intraperitoneal injection of infected blood
	Trichinella spiralis	Intragastric administration
Tumor cells	B16-F10 melanoma	Cells are injected intravenously
	PYB6 fibrosarcoma	Cells are injected subcutaneously

11.5 SPECIFIC CHEMICALS THAT ADVERSELY AFFECT THE IMMUNE SYSTEM

The number of drugs and chemicals associated with immunotoxicity in humans is extensive. As discussed in Section 11.2, immunotoxicity typically occurs as a hypersensitivity reaction, immune-suppression, or autoimmunity. Several agents commonly encountered in occupational settings are capable of producing contact, cell-mediated hypersensitivity, with common symptoms of rash, itching, scaling, and the appearance of redness and vesicles on the skin. Examples of these agents are shown in Table 11.3. The respiratory tract is also a common site of allergic symptoms from drug or chemical exposure. Inhalation of respiratory allergens can cause an immediate-type reaction (an *early-phase reaction*, occurring and waning rapidly) or a delayed-type reaction (sometimes called a *late-phase reaction*), which may appear 6–8 h later and require 12–24 h to resolve. Both reactions are IgE-mediated. Table 11.4 lists examples of common agents associated with respiratory allergy. *Occupational asthma* represents a special kind of inhalation disorder that is distinct from typical respiratory allergy. In general, a longer sensitization period is required, and symptoms may resemble an early-phase reaction, a late-phase reaction, or both. IgE may be responsible for some, but not all, of the manifestations of occupational asthma. In fact, the role of the immune system in occupational asthma may be different for asthma initiated or provoked by high-molecular-weight compounds, low-molecular-weight compounds, and irritants.

The potential for immunosuppression from occupational and environmental exposure to chemicals has been suggested by numerous *in vitro* studies and experiments in laboratory animals. Direct evidence for clinical immunosuppression following workplace or environmental exposures is extremely limited. However, there are many well-documented examples of the development or exacerbation of autoimmunity from chemical exposure. Most of these examples (shown in Table 11.5) are drugs, and for agents such as procainamide, up to 80% of patients treated chronically will develop increased levels of autoimmune antibodies. Many of these drugs produce signs and symptoms

TABLE 11.3 Examples of Agents that Produce Dermal Contact Sensitivity

Drugs	Resins
Benzocaine	Acrylic resins
Dibucaine	Epoxy resins
Thimerosal	Formaldehyde resins
Neomycin	Phenolic resins
Consumer products	Other industrial chemicals
Natural rubber latex (e.g., gloves, condoms)	Ethylenediamine
Preservatives (e.g., benzophenone, sorbic acid)	Paraphenylenediamine and other dyes
Metals	Antioxidants
Beryllium	Chlorinated hydrocarbons
Cadmium	Dinitrochlorobenzene
Chromates	Mercaptans
Gold	
Mercury	
Nickel	
Silver	
Zirconium	

TABLE 11.4 Examples of Agents that Produce Respiratory Allergy

Molds	Dusts and small particulates
Aspergillus	Coffee
Cladosporum	Enzymes
Hormodendrum	Flour
Penicillium	Mites
Rhizopus	Sawdust
Pollens (various)	Pet dander
	Cockroach proteins

TABLE 11.5 Examples of Agents Associated with Autoimmune Disease

Drugs
Acebutalol
Allopurinol
Alprenolol
Amiodarone
Ampicillin
Bleomycin
Captopril
Carbamazepine
Cephalosporin
Chlorpromazine
Chlorthalidone
Dapsone
Diphenylhydantoin
Ethosuximide
Fenoprofen
Iodine
Isoniazid
Lithium
Lovastatin
Mefenamic acid
Methyldopa
Minocycline
Nitrofurantoin
Penicillamine
Phenylbutazone
Propylthiouracil
Quinidine
Sulfonamides
Amino acids
L-Tryptophan
L-Canavanine
Environmental/industrial chemicals
Aromatic amines
Cadmium
Chlordane
Chorpyrifos
Chromium
Formaldehyde
Gold
Hydrazine
Mercury
Paraquat
Pentachlorophenol
Perchlorethylene
Silicon (silica)
Thallium
Trichloroethylene
Vinyl chloride

resembling SLE, while others produce autoimmune disease of the kidney, liver, thyroid, and other organs; scleroderma; or autoimmune hemolytic anemia. Evidence suggests that several environmental contaminants may also have the ability to either produce or worsen autoimmune disease, although the association with autoimmune disease is often less well-substantiated.

Some classes of chemicals or agents have, in particular, been associated with immunotoxic effects in humans. These are discussed briefly in the following text.

Metals

Metals have been associated with various types of hypersensitivity reactions. Beryllium, nickel, chromium, cadmium, silver, and zirconium have all been found to produce contact dermatitis. Nearly 10% of women and 2% of men have sensitivity to nickel and may develop rashes upon contact with nickel in jewelry, coins, and clothing fasteners. Sensitive individuals may also respond to chromium in tanned leather products. Metals are also associated with pulmonary hypersensitivity reactions and occupational asthma. One of the most serious of these diseases is *berylliosis*, a delayed hypersensitivity (type IV) reaction thought to result from beryllium acting as a hapten. Acutely, hypersensitivity to beryllium is manifested as pneumonitis and pulmonary edema. Chronically, workers exposed to beryllium develop a severe, debilitating granulomatous lung disease.

Studies in experimental animals have shown that metals such as lead, mercury, nickel, and cadmium are associated with activation of $CD4^+$ T cells or cause suppression of antibody responses and cell-mediated immunity, resulting in increased susceptibility to infection. There is some clinical and epidemiologic evidence that lead may decrease resistance to infectious disease, and the use of arsenic for medicinal purposes suggests that it, too, may have immunosuppressive effects. Arsenic was used in the early twentieth century to treat some inflammatory diseases, and currently appears to have some efficacy in treating leukemia. Also, patients treated with arsenicals were reported to have a relatively high incidence of the viral disease herpes zoster, suggesting some impairment of the immune system. More recent studies in mice have also indicated that nickel directly engages TLR4, a receptor in the innate immune system specific for bacterial lipopolysaccharide.

A number of studies have reported increased or unusual autoantibodies in association with exposure to some metals in the workplace, suggesting potential autoimmune toxicity. For example, there is evidence of immune complex glomerulonephritis in nephrotoxicity from cadmium and mercury. Iodine and lithium have been linked to autoimmune thyroid disease, and chromium and gold have been associated with SLE-like disease.

Polychlorinated Dibenzo(p)dioxins

Studies in rodents have shown that perinatal exposure to 2,3,7,8-tetrachlorodibenzodioxin (TCDD) appears to affect the developing thymus, leading to a persistent suppression of cellular immunity. The depression of T-cell function

from perinatal exposure appears to be greater and more persistent than when exposure occurs in adults. Molecular studies have shown that TCDD is a potent ligand for the cytosolic aryl hydrocarbon receptor (AhR). The effect appears to be tissue-specific. In the murine thymus, it can induce apoptosis (programmed death), while in other tissues, it can act as a tumor promoter, probably through actions on the tumor suppressor p53. The potential for TCDD immunotoxicity in humans is less clear. Individuals exposed to very high TCDD doses during an industrial explosion in Seveso, Italy in 1976 have not shown demonstrable loss in immune function. Studies of individuals exposed to TCDD chronically in Times Beach, Missouri have revealed a few differences from a control population in some parameters, but overall the observations do not suggest significantly altered immunocompetence. These studies have focused on humans exposed as adults to TCDD, and it is possible that perinatal exposure to TCDD may have more profound effects, as has been observed in laboratory animals. Increased antinuclear antibodies and immune complexes have been reported in blood of dioxin-exposed workers, but increases in clinical manifestations of autoimmunity have not been observed.

Dust and Particulates

A number of occupations involve inhalation exposure to high-molecular-weight organic molecules or particles containing these molecules. Examples include flour and wood dust; enzymes (e.g., from *Bacillus subtilis* and *Aspergillus niger* in the detergent industry); dusts from agricultural wastes; fungi and bacteria in moldy hay, feeds, and wood products; and dander, feces, papae, and other residue from insect and rodent pests. These high-molecular-weight substances are capable of producing an IgE-mediated, type I allergic reaction. This reaction can manifest itself as eye and upper respiratory tract congestion, occupational asthma, and hypersensitivity pneumonitis. Acute inhalation of dust from bacterial or animal origin has also been shown to produce a short-term flulike illness called "*organic dust toxic syndrome*." This is not a type I allergic reaction because no prior sensitization is required, nor are antigen-specific antibodies present during the illness. Inhalation of silica dusts both activates and damages alveolar macrophages. Activation of these macrophages can lead to pulmonary inflammation. Reported effects on lymphocyte responsiveness are somewhat conflicting, but suggest that immune function may be impaired. Inhalation exposure to crystalline silica dust, in addition to producing silicosis, has been associated with increased incidence of a variety of autoimmune diseases, including scleroderma, rheumatoid arthritis, and SLE. Generally, these associations have been observed in populations with high levels of silica exposure.

Pesticides

Dermal and pulmonary symptoms among workers handling pesticides are not uncommon, but most of these cases appear to be due to irritant rather than hypersensitivity reactions. Studies of workers exposed to pesticides have sometimes found changes in various specific immune parameters, but there is currently little evidence that host resistance is compromised in these individuals. A number of studies have found the presence of specific autoantibodies to be increased in populations exposed to pesticides, and some have linked pesticide exposure to autoimmune diseases such as lupus, rheumatoid arthritis, and immune complex glomerulonephritis. Methodological limitations and inconsistencies among studies limit the strength of evidence for autoimmune effects of pesticides in humans, however. Studies in animals suggest that some chlorinated pesticides may accelerate the development of autoimmunity, although no studies are yet available to assess whether this occurs in humans as well.

Solvents

Benzene is capable of producing bone-marrow hypoplasia and pancytopenia. Along with other formed elements of the blood, peripheral blood lymphocyte counts are diminished, leading to impaired immune function. Immunotoxic effects of benzene may extend beyond individuals experiencing bone-marrow toxicity from benzene, as humans exposed chronically to benzene have been observed to have diminished serum Igs and immune complement. Studies have indicated that metabolites of benzene adversely affect early transduction signals in lymphocytes.

Immune abnormalities, such as alterations in serum Ig concentrations, immunocyte counts, or immunocyte ratios have been observed in workers exposed to solvents, either individually or as mixtures. The significance of these findings is unclear, however, as no deficits in host resistance or other clinical immune effects have been demonstrated. Exposure to vinyl chloride has been linked to the development of scleroderma, and there is epidemiologic evidence of an association between chronic exposure to trichloroethylene in groundwater and lupus syndromes.

Miscellaneous Agents

In 1981, thousands of individuals in Spain were poisoned with cooking oil adulterated with rapeseed oil containing aniline. The symptoms that developed were called *toxic oil syndrome*, and included pneumonitis, rash, gastrointestinal distress, and marked eosinophilia. These patients developed autoantibodies and a connective tissue disorder characterized by myalgia, neuropathy, myopathy, and cutaneous manifestations. Hundreds of poisoned patients died, attributed primarily to impairment of respiratory musculature.

Acid anhydrides are used to produce a number of commercial products, including paints and epoxy coatings. On inhalation exposure, acid anhydrides can become haptens, binding to carrier proteins in the respiratory tract to elicit an immune response. After sensitization, subsequent exposure leads to asthma-like symptoms or to a reaction resembling hypersensitivity pneumonitis. Chronic exposure may lead to severe restrictive lung disease.

11.6 SUMMARY

A fully functioning immune system is vital for defense against pathogenic microorganisms and to prevent the emergence of cancerous cells. It is a complex system, requiring the cooperation of many types of cells. The immune system is capable of both specific and nonspecific responses to insults. Specific responses are elicited by macromolecules recognized by the body as being foreign, termed antigens. The presence of an antigen can trigger a humoral response (i.e., the production of antibodies that bind rather specifically to that molecule) or a cell-mediated response in which cells carrying the antigen on their surface are attacked by specialized immune cells (e.g., NK cells or cytotoxic lymphocytes).

Drugs and chemicals can produce adverse health effects by influencing the immune system in one of three ways:

1. *Causing a hypersensitivity reaction.* There are four basic types of hypersensitivity reactions (types I–IV), each with a different mechanism. Depending on the type of reaction, symptoms may be immediate or delayed, mild or severe, and involve different organs and tissues. Allergic reactions can cause considerable discomfort in the workplace, and some types (e.g., a severe type I reaction, or anaphylaxis) can be life-threatening.
2. *Suppressing the immune system.* Normal function of the immune system requires participation by many components, and disruption of any of these could conceivably result in impaired capability. If impairment is sufficient, the individual is at increased risk of infection and cancer. This has been clearly demonstrated by patients on immunosuppressive therapy (e.g., transplant patients) and in animal studies involving a variety of chemicals. Although there are few clear examples of immunosuppression from occupational or environmental exposure in humans, there is no reason to expect that this effect cannot occur under these circumstances as well.
3. *Causing or exacerbating autoimmune disease*. By producing a dysregulation of the immune system, drugs and chemicals are capable of causing the immune system to attack normal body constituents. This has been clearly demonstrated for several drugs, and a number of reports suggest that it may also occur from occupational and environmental exposures.

The potential for a chemical to produce immunotoxicity can be assessed through a variety of *in vivo* and *in vitro* tests. Most of these tests focus on effects on a very special aspect of the immune system. The immune system possesses considerable functional redundancy and extra capacity, and alterations (or "abnormalities") in one or a few parameters may not necessarily result in diminished overall functional of the immune system. Consequently, the results of these tests must be interpreted carefully.

SUGGESTED READING

Burleson GR, Dean JH, Munson AE. *Methods in Immunotoxicology*. New York: Wiley-Liss; 1995.

Descotes J. *Immunotoxicology of Drugs and Chemicals: An Experimental and Clinical Approach, Volume 1: Principles and Methods of Immunotoxicology*. New York: Elsevier Science; 2004.

Descotes J. Methods of evaluating immunotoxicity. Expert Opin Drug Metab Toxicol 2006;2:249–259.

Dietert RR, Luebke RW. *Immunotoxicity, Immune Dysfunction, and Chronic Disease*. New York: Humana Press; 2012.

Germolec DR, Kason M, Nyska A, Kuper CF, Portier C, Kommineni C, Johnson KA, Luster MI. The accuracy of extended histopathology to detect immunotoxic chemicals. Toxicol Sci 2004; 82:504–514.

Ladics GS. Use of SRBC antibody responses for immunotoxicity testing. Methods 2007;41:9–19.

U.S. Food and Drug Administration. *Guidance for Industry: Immunotoxicology Evaluation of Investigational New Drugs*. Rockville: USFDA; 2002.

12

TOXIC EFFECTS ON REPRODUCTION

MARY L. HIXON

That human exposure to environmental chemicals can cause reproductive harm has been known for centuries. Exposure to lead was recognized in Roman times as causing miscarriages and infertility in both men and women. Over the past several decades, reports of decreased semen quality and increased rates of developmental abnormalities of the male reproductive tract along with an increasing incidence of testicular cancer have focused attention on man-made chemicals as potential causative factors. Based on epidemiological findings, scientific research has concluded that these adverse reproductive trends have coincided with the dramatic increase in the production and exposure of both humans and wild-life to industrial and environmental chemicals. Consequently, a large number of natural and synthetic chemicals have now been studied and reported to disrupt the normal functioning of the reproductive system resulting in adverse effects in hormone-responsive target tissues and organs in both humans and animals.

This chapter will discuss adverse effects on reproduction in both men and women resulting from chemical exposure. Specific topics will include:

- The physiology of the reproductive systems of both men and women
- Toxicants that affect reproductive capability in men
- Toxicants that affect reproductive capability in women
- Methods to assess reproductive toxicity

12.1 GENERAL PHYSIOLOGY OF THE MALE AND FEMALE REPRODUCTIVE SYSTEM

Reproductive toxicology is the study of the adverse health effects of physical and chemical agents on the reproductive system in adult males and females. In addition, it includes the assessment of adverse reproductive effects observed in the embryo, fetus, neonate, and postnatal and pubertal organism following exposure to a chemical or physical agent. Therefore, the ability to determine the mechanism(s) of action for reproductive toxicants depends on a thorough understanding of the reproductive physiology of both the male and female reproductive systems and the potential sites of toxicant action (e.g., reproductive organ, tissue, cell type). The functions of the male and female reproductive organs or gonads are (i) to produce germ cells and (ii) to secrete sex hormones. The male reproductive organ or testes secrete primarily androgens (i.e., testosterone) along with small amounts of estrogen; and, the female reproductive organ or ovaries secrete mainly estrogen and small amounts of androgens. In addition, to estrogen and androgens, the ovaries secrete progesterone which prepares the uterus for implantation and pregnancy.

Male and female reproductive processes depend upon the proper functioning of the hypothalamic–pituitary–gonadal (HPG) axis. The hypothalamus functions as the interface between the central nervous system and the endocrine system. Neurons in the hypothalamus secrete a peptide known as gonadotropin-releasing hormone (GnRH), which bathes the cells of the anterior pituitary. These gonadotrope

Principles of Toxicology: Environmental and Industrial Applications, Third Edition. Edited by Stephen M. Roberts, Robert C. James and Phillip L. Williams.

cells respond to GnRH by secreting follicle-stimulating hormone (FSH) and luteinizing hormone (LH). FSH and LH then enter the bloodstream and stimulate the gonads to perform gametogenesis and hormone synthesis.

In the male, the HPG axis influences spermatogenesis, secondary sex characteristics, and sexual behavior. Spermatogenesis begins at puberty and is regulated by the steroid hormone testosterone, which is produced by the Leydig cells of the testis. The Leydig cells are located within the interstitium of the seminiferous tubules. As illustrated in Figure 12.1, the secretion of testosterone is regulated by LH, which stimulates androgen release in the male. An additional hormone which is important in spermatogenesis is FSH. FSH acts on Sertoli cells to promote the maturation of sperm. When testosterone is produced in adequate quantities, it produces a negative feedback effect which inhibits the production of GnRH, thus signaling to the hypothalamus to inhibit the synthesis of testosterone. When testosterone levels decline, GnRH is released and testosterone production is stimulated.

Similar to the male reproductive system, the female reproductive system is under control of GnRH released by the hypothalamus which stimulates secretion of LH and FSH by the anterior pituitary (Figure 12.1). The major functions of the ovary are the production of the female germ cell (oocyte) and the production of the female sex steroid hormone, estrogen. Secretion of LH and FSH stimulates the ovary to secrete both estrogen and progesterone. LH stimulates ovulation and the development of the corpus luteum; and, FSH promotes follicle development within the ovary allowing certain oocytes to mature and the follicle cells surrounding each oocyte to produce estrogen in preparation for fertilization. As with the male reproductive system, the female reproductive hormones participate in a complex negative regulatory feedback system that regulates release of these hormones in the female.

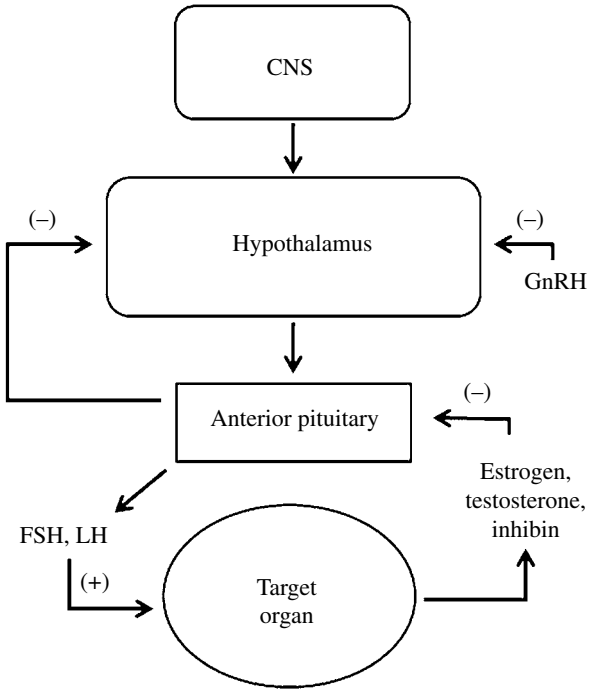


FIGURE 12.1 Hypothalamic–pituitary–gonadal axis. Endocrine feedback loops of the mammalian hypothalamic–pituitary–gonadal axis. Positive feedback is denoted by (+) and negative feedback is denoted by (–). FSH, follicle-stimulating hormone; LH, luteinizing hormone; GnRH, gonadotropin-releasing hormone.

The Male Reproductive System

The male reproductive system is composed of the testes, which serve as the site for gamete production and steroid hormone synthesis; the seminal vesicles and epididymes, which allow for the transport, maturation, and storage of sperm; and the penis, for the delivery of mature sperm to the female reproductive tract. The testes possess xenobiotic metabolizing enzymes and specific germ cell types differ in their metabolic function. Moreover, spermatogonia possess significant DNA repair enzymes and are fairly resistant to toxicants; whereas, actively dividing spermatogonia are much more susceptible to ionizing radiation and chemotherapeutic agents. All of these characteristics determine susceptibility of the testis to toxicants.

The process of spermatogenesis is a complex and delicate balancing act consisting of mitosis, meiosis, differentiation, and apoptosis. It involves the division and differentiation of germ cells to produce mature spermatozoa. The three stages of germ cell maturation include spermatocytogenesis, meiosis, and spermiogenesis. Tight junctions between adjacent Sertoli cells create two separate compartments within the seminiferous epithelium: a basal compartment below the tight junction and an adluminal compartment above the tight junctions. Spermatogenesis occurs in specialized thick-walled tubules within the testis called the seminiferous tubules. Initially, immature germ cells are located at the basement membrane or outside of the tubules, but as germ cells progress toward maturation, they migrate toward the tubular lumen. Spermatogonial stem cells undergo self-renewal within the testis and are responsible for the maintenance of spermatogenesis in the adult male. This proliferation is tightly regulated by endocrine, paracrine, and autocrine factors. As depicted in Figure 12.2. certain populations of spermatogonial stem cells will remain as stem cells; whereas, other stem cells will differentiate into spermatocytes. Primary spermatocytes undergo meiosis I to form two haploid secondary spermatocytes. These secondary spermatocytes undergo meiosis II to become four round spermatids. The spermatids remain connected by cytoplasmic bridges or syncytium (Figure 12.2). Spermiogenesis or the final stage of spermatogenesis allows for the maturation of the elongate

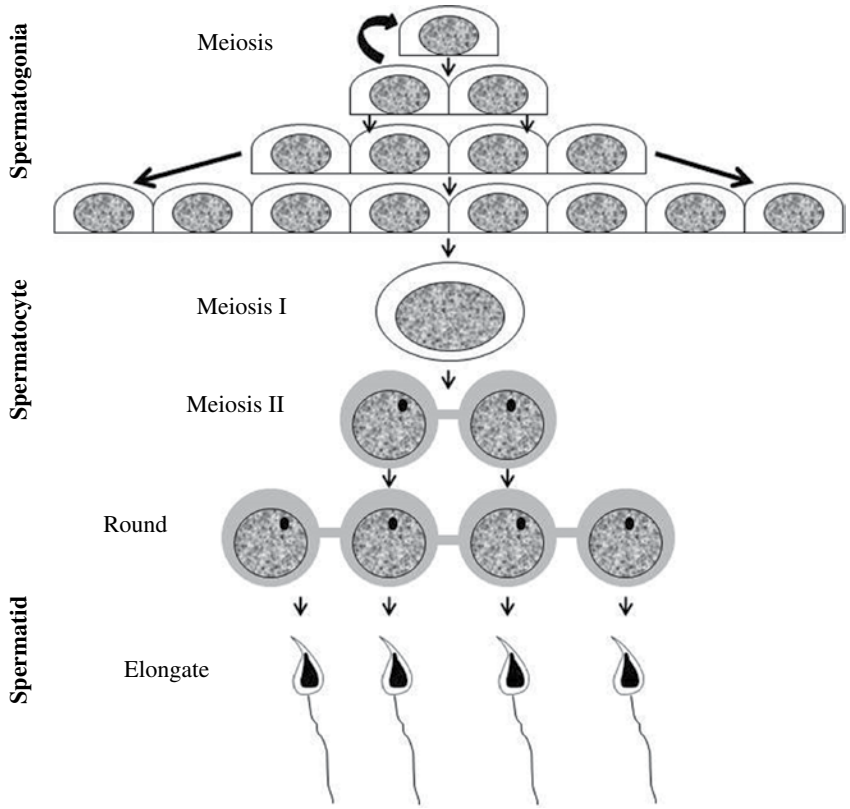


FIGURE 12.2 Spermatogenesis. The diagram highlights the mitotic spermatognia, the meiotic spermatocytes, and the haploid spermatids. The synctium or cytoplasmic bridges are depicted as gray bridges between the spermatocytes and round spermatids.

spermatids into sperm. In the human male, it takes approximately 64 days to produce a mature sperm. At its most productive, the average male can produce over 100 million sperms per day.

The maturation of the spermatids also involves the Sertoli cells, which are considered to be the support cells of the testis. They are highly specialized cells which are in direct contact with germ cells and provide both structural and functional support. The Sertoli cell functions in (i) the structural and nutritional support of the maturing germ cells; (ii) the movement of maturing germ cells from the basal to the adluminal compartments of the seminiferous epithelium; (iii) the spermiation of mature spermatids; (iv) the secretion of luminal fluids; (v) the removal of waste products; and (vi) inhibin biosynthesis. A single Sertoli cell may provide support for 20–30 germ cells at various spermatogenic stages and the Sertoli cell accomplishes this support by communicating to the germ cells through specialized gap junctions. Sperm production in the testis directly correlates with the number of healthy Sertoli cells in the testis, and any toxicant that disturbs Sertoli cell function may adversely alter spermatogenesis. It is important to note that proper Sertoli cell function is dependent upon hormone signaling. Testosterone production via Leydig cells promotes healthy Sertoli cell function; and, FSH, which acts directly on Sertoli cells, also promotes healthy spermatogenesis. Therefore, toxicants that inhibit testosterone synthesis in Leydig cells and/or testosterone-mediated functions of Sertoli cells may impair spermatogenesis. Figure 12.3 illustrates the tight junctions which form between Sertoli cells to provide a barrier known as the blood–testis barrier, which keeps the developing spermatozoa separate from the immune system. The ability of a toxicant to target specific germ cell type(s) depends on its ability to penetrate this barrier.

The Female Reproductive System

The female reproductive tract consists of the ovaries; oviduct or fallopian tubes; the uterus; cervix and vagina; and the mammary glands, which provide nutritional support to the offspring during the neonatal period. In addition to gamete production, the female reproductive system is the site for fertilization and pregnancy and provides nutritional support for the offspring. Puberty is the point at which reproduction becomes possible in the female and is signaled by the first menstruation. Menopause or ovarian failure in women is associated with the cessation of ovarian cyclicity. It is becoming more appreciated in the scientific literature that in utero and/or postnatal exposure(s) to chemicals has the potential to influence not only the timing of puberty in young

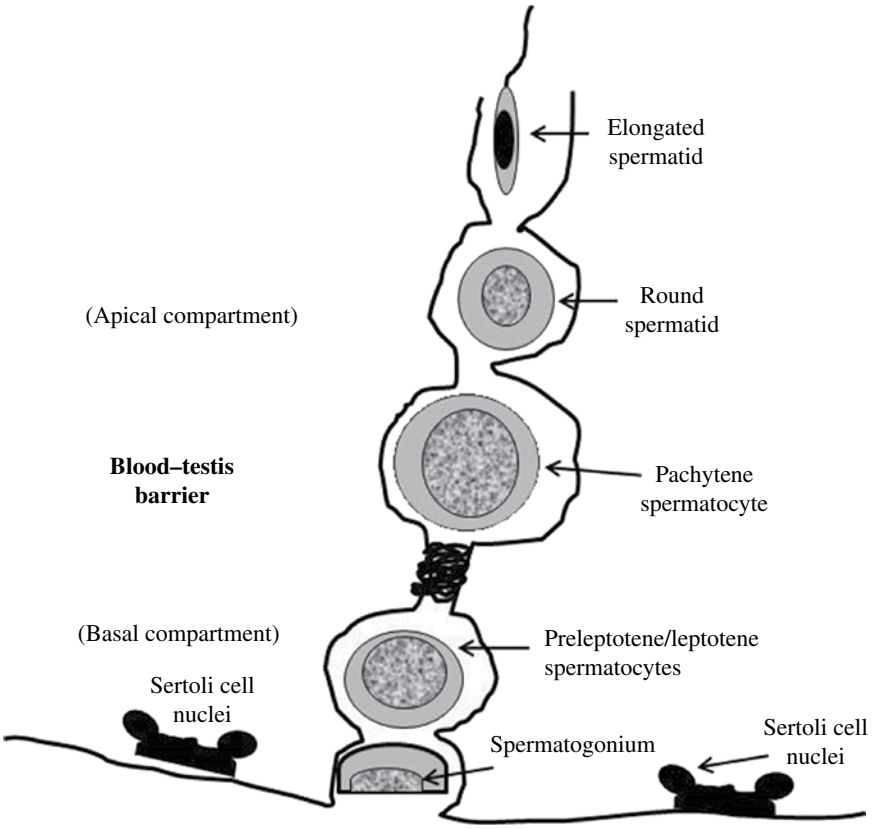


FIGURE 12.3 The blood–testis barrier (BTB). The blood–testis barrier separates the seminiferous epithelium into basal and apical compartments. In the testis, tight junctions between Sertoli cells are an important structural element of the BTB restricting the movement of water, solutes, and immune cells from the circulation to the adluminal compartment of the seminiferous tubule, which creates a unique microenvironment for spermatogenesis. The developing spermatocytes pass through the blood–testis barrier as they move toward the lumen of the seminiferous tubule. The BTB undergoes constant restructuring without compromising the integrity or function of the barrier.

girls; but also, the onset of peri-menopause and, consequently, menopause. Even a small shift in the female reproductive life span can have deleterious consequences to female reproductive health and fertility.

The development and maturation of the oocyte occurs within the ovarian follicles, and ovulation absolutely requires the proper development of the ovarian follicles. The most immature stage of follicular development is the primordial follicle (Figure 12.4). These follicles first appear in the ovary of the developing female fetus. The primordial follicle is then activated and develops into a primary follicle. As follicular development continues, there is a rapid proliferation of granulosa cells surrounding the oocyte and the acquisition of a layer of theca interna cells surrounding the granulosa cells. The primary follicle develops into a secondary follicle upon the development of multiple layers of granulosa cells (Figure 12.4). Eventually, an antrum or fluid-filled space develops within the granulosa cell layers. This antral follicle continues to grow and at its most mature stage is referred to as a Graaffian or preovulatory follicle (Figure 12.4).

During the reproductive life span of the female, primordial oocytes are continuously recruited to mature with every menstrual cycle. The process of recruitment and maturation to ovulation is termed folliculogenesis. Over the lifetime of the female, it is estimated that only 400–500 oocytes are ovulated. The remaining follicles will die by atresia. By age 50 in the human female, the pool of oocytes is estimated to be less than 1000. This depletion of oocytes is referred to as menopause.

Following ovulation, the remaining granulosa cells in the follicle undergo terminal differentiation to form the corpus luteum. The corpus luteum secretes large amounts of progesterone, estradiol, and inhibins. Inhibins function to reduce secretion of gondaotropins from the hypothalamic pituitary axis via negative feedback; thus halting maturation of other ovarian follicles. Inhibins also increase the secretory activity of the uterine endothelium in preparation for implantation of a fertilized oocyte. If fertilization does not occur, the corpus luteum regresses, resulting in the rapid decreases of progesterone levels, shedding of the uterine lining, and resumption of the ovarian cycle.

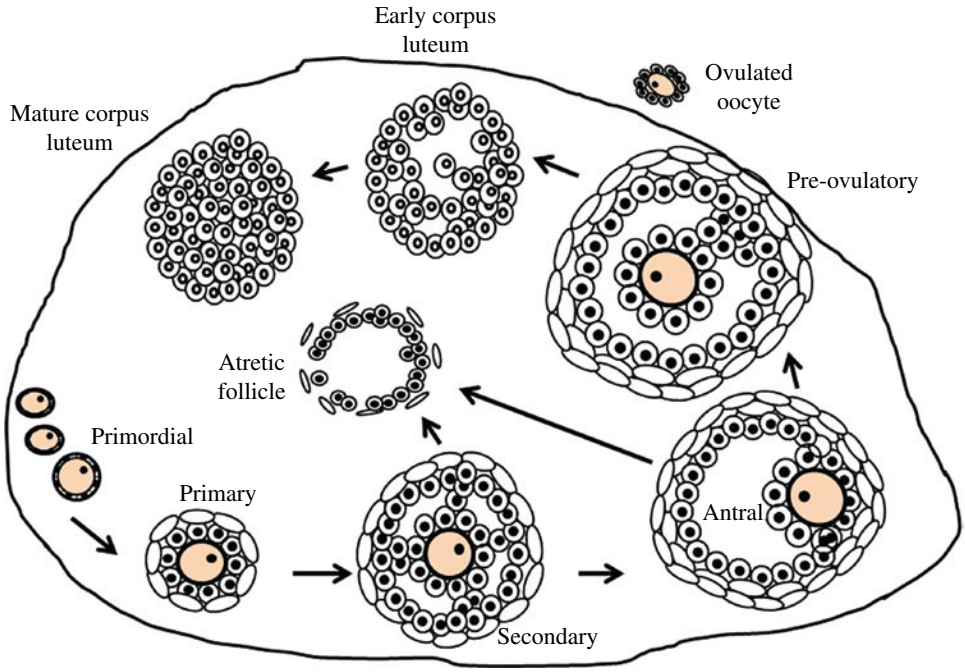


FIGURE 12.4 The ovary. The development of ovarian follicles is illustrated in the diagram. Primordial follicles are activated to grow and develop from primary through secondary to antral stages of development. Most follicles degenerate before reaching the ovulatory stage. Follicles that are ovulated then luteinize to become corpora lutea, which support pregnancy if fertilization occurs.

12.2 GENERAL METHODS TO ANALYZE REPRODUCTIVE TOXICANTS

The level of complexity of the male and female reproductive system in determining a link between an adverse reproductive effect and a chemical poses a significant challenge to the reproductive toxicologist. Traditional reproductive toxicity studies are a major source of data on the effects of potential reproductive toxicants in both men and women. These types of studies include structural and functional alterations that may affect reproductive competency such as fertility, parturition, and lactation. Studies to examine maternal and paternal behaviors also provide information on gonadal function, estrous cyclicity, mating behavior, conception, parturition, lactation and weaning, and the growth and development of the offspring. The purpose of traditional reproductive toxicology studies is to determine a no-observable-adverse-effect level or NOAEL for the test chemical in question. It should be noted that these types of studies do not provide information on the maternal endocrine environment or the levels of parent compound or metabolites in the milk or fetus of the pup.

Endpoints for identifying reproductive toxicants include measures of primary reproductive function and behavior. Table 12.1 lists some of the reproductive parameters examined in traditional *development and reproductive toxicology* studies referred to as **DART**. These categories include both male- and female-specific endpoints and couple-mediated

TABLE 12.1 Endpoints for Reproductive Studies

Reproductive Endpoint	Parameters Examined
Male-specific	Body weight
	Mating behavior
	Reproductive organ weights
	Histopathology
	Sperm transport, maturation, and storage
	Seminal fluid production
	Testosterone secretion and production
Female-specific	Body weight
	Mating Behavior
	Reproductive organ weights
	Estrous cycle onset and length
	Histopathology
	Ovulation
	Estradiol secretion and production
	Gestational length
	Corpora lutea number
	Pre- and postimplantation loss
	Parturition
	Lactation and nursing behavior
Couple-mediated	Mating behavior
	Pregnancy rate
	Litter size
	Birth weight
	Offspring survival
	Sex ratio
	External malformations

endpoints. Couple-mediated endpoints are defined as those endpoints which can have a contributing role to toxicity when both couples are exposed to a chemical agent. Alterations in these endpoints may be the result of direct or indirect toxicity to the male and/or female. In addition, the evaluation of dose response curves for target organ effects provides an estimation of the dose levels of specific target organ effects which predict the potential adverse reproductive effects of human exposure(s).

For the male, monitoring body weight during exposure to a chemical agent provides an index of overall general health status and is vitally important for the interpretation of reproductive effects. Testis weight varies only modestly within a test species and changes in absolute testis weight are a good indicator of gonadal injury. However, changes may only be noticed at doses higher than those required to induce changes in other reproductive indices. Prostate and seminal vesicle weights are androgen-dependent and therefore may reflect changes in endocrine status or testicular function.

Histopathology is another common endpoint in determining both male and female reproductive toxicity. Organs that are often evaluated in male reproductive toxicity evaluations include the testes, epididymides, prostate gland, seminal vesicles including the coagulating glands, and pituitary. Histopathology can provide a fairly sensitive indicator of damage with short-term dosing providing information on both the target cell(s) and the extent of toxicity of the chemical agent being tested. Due to the differential sensitivity of germ cells at various stages of maturation, the effects of a toxicant on spermatogenesis may be either immediate or delayed. For example a toxicant that targets mature spermatozoa in the epididymis will result in immediate infertility; whereas, a toxicant that targets primary spermatocytes will result in delayed infertility. In addition, germ cell death, germ cell sloughing, multinucleated germ cells, Sertoli cell vacuolization, altered seminiferous tubule diameter, and testicular atrophy are endpoints utilized to determine testicular toxicity.

For the female, the monitoring of body weight during exposure to a potential reproductive toxicant also provides an index of overall health of the female; and, as with the male, the female reproductive organ weights, ovaries and uterus, are also examined macroscopically and weighed. Toxicants which target the ovary are likely to affect oocyte number, follicle development, ovulation, cyclicity, fertility, and/or pregnancy maintenance. Histology and ovarian morphometry are generally utilized to evaluate dose and the follicle(s) targeted. Histology is generally utilized as part of the framework of a female reproductive toxicity evaluation. Serial sections of the ovary are analyzed with generally every 10th section utilized for analysis. Properly conducted follicle counts can supplement qualitative ovarian assessment to characterize ovarian toxicants, to understand a toxicants site(s) of action, and to assess primordial follicle integrity when ovarian lesions are subtle.

Corpora lutea number is also utilized as a quantitative assessment in female reproductive toxicology. Following ovulation, a corpora lutea is formed at the site of ovulation. Corpora lutea are formed due to the differentiation of granulosa cells and can indicate that a chemical is a reproductive toxicant if the numbers of corpora lutea are reduced. Preimplantation loss can also be utilized as a method for determining toxicity. However, it must be noted that preimplantation loss can be affected by the number of eggs ovulated, fertilized, implanted, and also the receptivity of the uterus.

The Mammary Gland

Besides the potential reproductive effects of toxicants on ovarian physiology, impaired lactation of the mammary gland by toxicants may affect changes in fertility. Development of the breast is a major event at the onset of puberty. Estrogens are the primary hormones responsible for mammary gland development and growth. After childbirth or parturition, milk secretion is initiated by suckling of the nipple. Normal milk stimulation and secretion requires multiple hormones including prolactin, adrenocorticotropic hormone (ACTH), insulin, growth hormone, and thyroid hormone. Therefore, toxicants which block the synthesis and/or actions of estradiol and progesterone may compromise the development and differentiation of the mammary gland during puberty and pregnancy. Studies in rodents suggest that gestational exposure to the chemical perfluorooctanoic acid (PFOA) results in delays in mammary gland development and/or lactational differentiation in both dams and their offspring.

New Technologies to Assess Reproductive Toxicants

The traditional animal study designs for assessing both reproductive and developmental toxicity cannot accommodate the evaluation of large numbers of chemicals and require the development of alternative technologies. The development of high-throughput techniques to analyze the genome, transcriptome, proteome, and and, more recently, the metabolome are revolutionizing the way scientists and health professionals study not only biology; but also, reproductive toxicology. Genomic and proteomic tools have been developed and applied to study changes in gene expression and cellular proteins in response to a variety of pharmaceutical and environmental chemicals in a variety of organisms. The metabolome represents a comprehensive data set that defines the small molecules present in a cell, tissue, whole organism, or biological fluid. From this information, it is now possible to establish relationships between metabolite levels and the cellular response of an organism to chemical

and/or nutritional stimuli that may be associated with a particular adverse effect or disease state.

Two of the most widely used platforms for metabolomic analyses are nuclear magnetic resonance (NMR) and mass spectrometry (MS). NMR technology has been used to detect changes in metabolites from data sets of hundreds of endogenous metabolites following exposure to drugs or environmental toxicants. NMR studies have also demonstrated metabolic changes in developing organisms as well as metabolic changes resulting in human disease. Mass spectrometry (MS) has emerged as a powerful tool for metabolomic analysis. This platform coupled with chromatography offers a powerful analytical tool with high sensitivity and the ability to determine chemical structures. The future challenge for reproductive toxicologists in both the industrial/pharmaceutical and the environmental setting is the integration of high-throughput technologies such as transcriptomics, proteomics, and metabolomics for prioritizing the many thousands of environmental chemicals with little to no hazard information.

12.3 REPRODUCTIVE TOXICANTS

Male Reproductive Toxicants

As previously stated, the human testis is a known target organ for toxicant-induced testicular injury resulting from exposure to therapeutic agents, chemical agents, and radiation. This susceptibility is due, in part, to the high rates of proliferation, differentiation, and metabolic activity associated with the production of large quantities of mature sperm. Within the testis, the three main target cells for toxicants that disrupt spermatogenesis are the Leydig, Sertoli, and germ cells. Numerous chemicals target various cell type(s) resulting in germ cell apoptosis and spermatogenic failure. The next section details a few examples of well-studied male reproductive toxicants from both animal and human studies. It is not intended to be an exhaustive list.

Spermatogonia and Spermatocytes The toxic effects of ionizing irradiation on spermatogenesis have been extensively studied. Germ cell death occurs primarily by apoptosis and results from oxidative damage, DNA single-strand breaks, and DNA double-strand breaks. As a consequence, actively dividing spermatogonia are the most susceptible germ cell population followed by spermatocytes. Methoxyacetic acid (MAA) is the active biological oxidation product of the industrial solvent ethylene glycol monomethyl ether (EGME). EGME causes acute toxicity in several species including humans. MAA primarily affects tissues with rapidly dividing cells and high rates of energy metabolism. Testicular toxicity, one of the most prominent consequences of EGME and MAA exposure, results from apoptosis of pachytene spermatocytes and is associated with changes in the expression of androgen responsive genes and oxidative stress.

Germ cell toxicants have the potential to cause both genetic and/or epigenetic changes that can be passed to subsequent generations. Genotoxic agents directly target the DNA by causing single-strand and double-strand breaks; whereas, epigenetic toxicants do not alter the DNA sequence but rather cause changes in methylation patterns. In animal studies, testicular toxicants that have been shown to induce epigenetic changes in male germ cells are the fungicide, vinclozolin, and the pesticide, methoxychlor (MXC). Besides epigenetic changes associated with MXC, perinatal and juvenile oral exposure to MXC in rodents has been reported to alter the spermatogenic potential of adult males by reducing the number of Sertoli cells within the testis.

Spermatids In men, cigarette smoking reduces sperm production and increases oxidative stress resulting in damage to the DNA. Epidemiological studies have shown that spermatozoa from smokers exhibit reduced fertilizing capacity, and embryos display lower implantation rates. Further studies on infertility patients have demonstrated that males with idiopathic infertility generate a higher percentage of sperm DNA fragmentation attributed to increased oxidative stress and insufficient scavenging of antioxidant enzymes in the seminal fluid.

Besides cigarette smoking, epidemiological studies have correlated exposure to pesticides with decreases in male fertility. Carbaryl is an extremely important carbamate insecticide and it has been used for over 30 years to control a wide range of pests, particularly in developing countries. Due to it widespread use, human carbaryl exposures from pesticide manufacturing, crop dusting, and daily life are common. Epidemiologic and occupational studies have found that carbaryl exposure correlates with adverse reproductive outcomes such as infertility, pregnancy loss, and stillbirth. However, the potential mechanisms of these toxic effects are not entirely clear. Studies have suggested that there is an important relationship between occupational carbaryl exposure and sperm genotoxic effects. Moreover, an elevated level of sperm morphologic defects has been detected in spermatozoa of carbaryl-exposed workers compared with control populations.

Moderately reduced sperm quality has also been suspected after occupational exposure to numerous other chemicals such as ethylene dibromide, chlordecone, and certain glycol ethers found in paint, glue, printing inks, and antifreeze solutions. Specific organic solvents such as styrene found in plastic coatings and chloroprene used in the production of plastics have been associated with poor semen quality. Sperm maturation may be impaired by alpha-chlorohydrin, antiandrogens, or sulphasalazine, and sperm motility in the

female genital tract can be inhibited by vaginal spermicides such as nonoxynol-9 and by propranolol.

The Sertoli Cell Phthalates, such as di-(2-ethylhexyl) phthalate, are ubiquitous environmental toxicants capable of producing testicular atrophy in laboratory animals. Mono-(2-ethylhexyl) phthalate (MEHP), the active metabolite of di-(2-ethylhexyl) phthalate, targets peri-pubertal Sertoli cells, resulting in a rapid induction of testicular germ cell apoptosis, with spermatocytes as the most sensitive population. One proposed mechanism of action for MEHP-induced testicular injury involves the up-regulation of FasL mRNA on Sertoli cells and the subsequent up-regulation of Fas mRNA on germ cells. This up-regulation leads to germ cell apoptosis and germ cell sloughing in which immature germ cells become detached from the seminiferous epithelium.

The solvent *n*-hexane is a Sertoli cell toxicant. Its testicular toxicity is due to the ability of its metabolite 2,5-hexandione to target microtubules. Microtubules are necessary for the formation of the Sertoli cell scaffolding which is the structural support of the Sertoli cell. Toxicants such as *n*-hexane and colchicine that target the microtubules, lead to germ cell sloughing and infertility. A chemical widely used in the manufacture of explosives, 1,3-dinitrobenzene (DNB), also targets Sertoli cells leading to germ cell apoptosis within the testis.

The Leydig Cell Leydig cells are the primary steroidogenic cell of the testis and they are located in the interstitial compartment of the testis. Leydig cell toxicity is primarily manifested by decreased testosterone synthesis, which leads to decreased testis and seminal vesicle weights and declining sperm production. One classic Leydig cell toxicant is ethane-1,2-dimethanesulfonate (EDS), a cytotoxic alkylating agent that has been shown to cause Leydig cell ablation, a rapid decrease in testosterone levels, and a characteristic pattern of germ cell loss in rats. This characteristic pattern of germ cell loss has been investigated using a variety of hormonally disruptive techniques and agents and is manifested as apoptosis of specific developmental stages of germ cells; over time, this stage-specific germ cell apoptosis results in the depletion of maturing germ cells from the seminiferous epithelium. The underlying molecular pathways affected in this hormone insufficiency response are the androgen-dependent transcriptional programs of testicular somatic cells.

The Blood–Testis Barrier Mature Sertoli cells create a barrier of tight junctions, also known as the "blood–testis barrier" (BTB) (Figure 12.3). The formation of the blood–testis barrier separates the testis into basal and apical compartments, and it also allows for a fluid-filled lumen to form. When the BTB is dysfunctional, germ cell differentiation and development are arrested. For example, cadmium (Cd), a major industrial pollutant and an endocrine disruptor, can cause significant damage to the testis by disrupting the blood–testis barrier. Occupational exposure to Cd usually takes place during mining or manufacturing of batteries and pigments. The testis is sensitive to Cd toxicity. Since the 1950s, studies have shown that *in vivo* acute exposure to Cd in animal models results in BTB disruption, germ cell loss, testicular edema, hemorrhage, necrosis, and sterility in several mammalian species. Recent studies have also associated reduced male fertility (e.g., reduced sperm count and poor semen quality) in men exposed to Cd.

Dibromochloropropane: A Case Study Dibromochloropropane (DBCP), a brominated organochlorine nematocide, gained popularity in the mid-1950s, because it was less acutely toxic than earlier soil fumigants to plants. By 1975, 25 million pounds of DBCP were being produced in the United States per year. Animal studies in the early 1960s revealed hepatic and renal effects and testicular atrophy in DBCP-treated rats; and testicular atrophy was noted at the lowest exposure level tested (5 ppm). In 1977, employees in a California pesticide formulation plant were discovered to be infertile. Further investigations documented azoospermia and oligospermia among many of the workers handling DBCP. These findings were duplicated in other DBCP production plants. In the United States, the Occupational Safety and Health Administration (OSHA) and the Environmental Protection Agency (EPA) rapidly restricted the use of DBCP. Before 1977, no such chemical in workplace concentrations had been shown to produce infertility and sterility in otherwise healthy men. The precise mechanism of action of DBCP is not completely understood, but a correlation between induced DNA damage and testicular necrosis suggests that DNA damage is the initiating event. Other short-chained halogenated alkanes are widely used in industry as chemical intermediates, extraction solvents, degreasing compounds, copolymer cross-linking agents, and pesticides. Some of these chemicals have also been shown to exhibit toxic effects on the testis; however, to date, DBCP is the most striking occupational testicular toxicant in men.

Estrogenic Chemicals The potential for environmental estrogenic exposure to estrogens in the male is real. In males, estrogens block spermatogenesis by suppressing LH and FSH, which in turn suppress testosterone secretion. The accessory sex glands of the male reproductive system, the seminal vesicles, the bulbourethral gland, and the prostate, are highly dependent on androgen production. Due to the fact that these tissues rely on testicular androgen synthesis for maintenance and function, changes in their weight serve as effective indicators of insufficient steroidogenesis.

Numerous drugs and environmental chemicals have been shown to affect male reproduction through endocrine-related modes of action. For example, the antihypertensive drug propanolol has been shown to decrease relative testis,

epididymal, ventral prostate, and seminal vesicle weights attributed to a significant decrease in testosterone, LH, and FSH. Exposure to the pesticide chlordecone resulted in men occupationally exposed to exhibit reduced fertility. Besides chlordecone, other pesticides such as mirex, DDT, the carbamates, and some of the PCBs are all reported to disrupt male fertility. The metabolites of DDT are weakly estrogenic as well as the polychlorinated biphenyls (PCBs). Polyaromatic hydrocarbons also possess weak estrogenic activity. Besides environmental and industrial chemicals, there are naturally occurring plants with estrogenic activity. For example, soy plants which contain genistein and zearalenone, a naturally occurring estrogenic mycotoxin have been shown to affect the male reproductive system.

Female Reproductive Toxicants

As with the male, reproductive function in women can be compromised by exposure to environmental chemicals. Chemicals that target the ovary can therefore have a significant effect on fertility, menstrual (estrous) cyclicity, and the timing of puberty and menopause. Ovarian toxicity is not simply confined to one cell type. The ovary is a complex structure composed of multiple cell types (e.g., granulosa, theca, and interstitial cells) and follicles in varying developmental stages (e.g., primordial, primary, secondary, antral, and corpora lutea); and studies indicate that various toxicants target a specific cell type(s) within a specific follicle(s). The next section details a few examples of well-studied female reproductive toxicants from both animal and human studies. It is not intended to be an exhaustive list.

Pesticides For women working in agriculture, exposure to pesticides is unavoidable. Epidemiology studies have found a positive correlation between organochlorine residues in maternal and cord blood placenta, and fetal tissues; and, an increase in the incidence of spontaneous abortion, premature delivery, and stillbirth in women. Moreover, chronic exposure to organophosphates has been associated with abnormal menstruation, amenorrhea, and early menopause.

Cigarette Smoke Epidemiological studies have demonstrated that cigarette smoking is associated with adverse reproductive outcomes in human females. Many toxicants present in cigarette smoke, such as benzoapyrene (B[a]P), dimethyl-benzenanthracene (DMBA), and 3-methylcholanthrene (3-MC) may be central to the documented adverse effects of cigarette smoke on follicular development and subsequent infertility. Experimental evidence indicates that B[a]P reaches the follicular fluid of the ovary and it is found at much higher levels in women who smoke than those women who do not smoke. In rodent models, exposure to all three of these polycyclic aromatic hydrocarbons has been shown to lead to the depletion of oocytes and the induction of ovarian tumors. Some of these reproductive toxicants must be metabolized in order to produce toxicity; and, activation can occur in the ovary as cytochrome P450 enzymes and other metabolizing enzymes are found in ovarian tissue.

4-Vinylcyclohexene The occupational chemical 4-vinylcyclohexene (VCH) has been shown to cause destruction of small pre-antral follicles in the ovaries of rodents. The monoepoxide metabolites, 1,2-VCH epoxide, 7,8-VCH epoxide, and the diepoxide, VCD, generated by the parent compound have been shown to cause pre-antral follicle loss in rodent models. Chemicals that destroy small pre-antral follicles are of serious concern to women of reproductive age because exposure can result in premature ovarian failure and consequently early menopause.

Ionizing Radiation Similar to the male, rapidly dividing germ cells are targeted in the female by radiation. Animal studies have shown that prepubertal mice exposed to irradiation leads to the destruction of primordial and primary follicles within the ovary. This is caused by the apoptosis of both granulosa cells and oocytes.

Cyclophosphamide Besides irradiation, cyclophosphamide is an alkylating chemotherapeutic agent associated with ovarian failure. The active metabolite of cyclophosphamide (CPA) in the ovary is phosphoramide mustard. In mice, CPA has been shown to significantly reduce the number of both primordial and antral follicles resulting in infertility. More recent studies have indicated the detection of DNA damage in oocytes of small ovarian follicles following phosphoramide mustard exposures of cultured rodent ovaries *in vitro* at levels that do not destroy the follicles. As with irradiation, women receiving cyclophosphamide as a chemotherapeutic agent may have reduced fertility as a consequence of the cancer regimen.

REFERENCES AND SUGGESTED READING

Anway MD, Memon MA, Uzumcu M, Skinner MK. Transgenerational effect of the endocrine disruptor vinclozolin on male spermatogenesis. J Androl 2006;27(6):868–879.

Bagchi G, Waxman DJ. Toxicity of ethylene glycol monomethyl ether: impact on testicular gene expression. Int J Androl 2008; 31(2):269–274.

Bagchi G, Zhang Y, Stanley KA, Waxman DJ. Complex modulation of androgen responsive gene expression by methoxyacetic acid. Reprod Biol Endocrinol 2011;9:42.

Baranski B. Effects of the workplace on fertility and related reproductive outcomes. Environ Health Perspect 1993;101(Suppl 2):81–90.

Benoff S, Jacob A, Hurley I. Male infertility and environmental exposure to lead and cadmium. Hum Reprod Update 2000;6:107–121.

Boekelheide K. Mechanisms of toxic damage to spermatogenesis. J Natl Cancer Inst Monogr 2005;(34):6–8.

Buck Louis GM, Sundaram R, Schisterman EF, Sweeney AM, Lynch CD, Gore-Langton RE, Maisog J, Kim S, Chen Z, Barr DB. Persistent environmental pollutants and couple fecundity: the LIFE study. Environ Health Perspect 2013;121(2):231–236.

Cai L, Hales BF, Robaire B. Induction of apoptosis in the germ cells of adult male rats after exposure to cyclophosphamide. Biol Reprod 1997;56:1490–1497.

Chevrier C, Warembourg C, Gaudreau E, Monfort C, Le Blanc A, Guldner L, Cordier S. Organochlorine pesticides, polychlorinated biphenyls, seafood consumption, and time-to-pregnancy. Epidemiology 2013;24(2):251–260.

Craig ZR, Wang W, Flaws JA. Endocrine-disrupting chemicals in ovarian function: effects on steroidogenesis, metabolism and nuclear receptor signaling. Reproduction 2011;142(5):633–646.

Creasy DM. In: *Histopathology of the male reproductive system II—evaluation*. Chapin RE, Boekelheide K, editors. New York: John Wiley & Sons, Inc.; 2002.

Davis JR, Brownson RC, Garcia R. Family pesticide use in the home, garden, orchard, and yard. Arch Environ Contam Toxicol 1992;22:260–266.

Devine PJ, Hoyer PB, Keating AF. Current methods in investigating the development of the female reproductive system. Methods Mol Biol 2009;550:137–157.

Devine PJ, Payne CM, McCuskey MK, Hoyer PB. Ultrastructural evaluation of oocytes during atresia in rat ovarian follicles. Biol Reprod 2000;63(5):1245–1252.

Devine PJ, Perreault SD, Luderer U. Roles of reactive oxygen species and antioxidants in ovarian toxicity. Biol Reprod 2012;86(2):27.

Devine PJ, Rajapaksa KS, Hoyer PB. In vitro ovarian tissue and organ culture: a review. Front Biosci 2002;7:d1979–d1989.

Dym M, Clermont Y. Role of spermatogonia in the repair of the seminiferous epithelium following x-irradiation of the rat testis. Am J Anat 1970;128:265–282.

Elshal MF, El-Sayed IH, Elsaied MA, El-Masry SA, Kumosani TA. Sperm head defects and disturbances in spermatozoal chromatin and DNA integrities in idiopathic infertile subjects: association with cigarette smoking. Clin Biochem 2009;42(7–8):589–594.

Feng Y, Fang X, Shi Z, Xu M, Dai J. Effects of PFNA exposure on expression of junction-associated molecules and secretory function in rat Sertoli cells. Reprod Toxicol 2010;30(3): 429–437.

Fenton SE. The mammary gland: a tissue sensitive to environmental exposures. Rev Environ Health 2009;24(4):319–325.

Hayes AW, editor. *Principles and Methods of Toxicology*. 5th ed. Philadelphia: Taylor & Francis; 2008. p 1605–1711.

Hoyer PB. Damage to ovarian development and function. Cell Tissue Res 2005;322(1):99–106.

Hoyer PB, Devine PJ, Hu X, Thompson KE, Sipes IG. Ovarian toxicity of 4-vinylcyclohexene diepoxide: a mechanistic model. Toxicol Pathol 2001;29(1):91–99.

Johnson L, Staub C, Silge RL, Harris MW, Chapin RE. The pesticide methoxychlor given orally during the perinatal/juvenile period, reduced the spermatogenic potential of males as adults by reducing their Sertoli cell number. Reprod Nutr Dev 2002;42(6):573–580.

Juhler RK, Larsen SB, Meyer O, Jensen ND, Spano M, Giwercman A, Bonde JP. Human semen quality in relation to dietary pesticide exposure and organic diet. Arch Environ Contam Toxicol 1999;37:415–423.

Knudsen TB, Kavlock RJ, Daston GP, Stedman D, Hixon M, Kim JH. Developmental toxicity testing for safety assessment: new approaches and technologies. Birth Defects Res B Dev Reprod Toxicol 2011. DOI:10.1002/bdrb.20315.

Lag M, Soderlund EJ, Omichinski JG, Brunborg G, Holme JA, Dahl JE, Nelson SD, Dybing E. Effect of bromine and chlorine positioning in the induction of renal and testicular toxicity by halogenated propanes. Chem Res Toxicol 1991;4:528–534.

Li L, Heindel J. Sertoli cell toxicants. In: Korach K, editor. *Reproductive and Developmental Toxicology*. New York: Marcel Dekker; 1998.

Mark-Kappeler CJ, Hoyer PB, Devine PJ. Xenobiotic effects on ovarian preantral follicles. Biol Reprod 2011;85(5):871–883.

Mayer LP, Pearsall NA, Christian PJ, Devine PJ, Payne CM, McCuskey MK, Marion SL, Sipes IG, Hoyer PB. Long-term effects of ovarian follicular depletion in rats by 4-vinylcyclohexene diepoxide. Reprod Toxicol 2002;16(6):775–781.

Meistrich ML. Male gonadal toxicity. Pediatr Blood Cancer 2009;53(2):261–266.

Meistrich ML, Wilson G, Shuttlesworth GA, Porter KL. Dibromochloropropane inhibits spermatogonial development in rats. Reprod Toxicol 2003;17(3):263–271.

Miller KP, Borgeest C, Greenfeld C, Tomic D, Flaws JA. In utero effects of chemicals on reproductive tissues in females. Toxicol Appl Pharmacol 2004;198(2):111–131.

Miller MG. Environmental metabolomics: a SWOT analysis (strengths, weaknesses, opportunities, and threats). J Proteome Res 2007;6(2):540–545.

Moffit JS, Bryant BH, Hall SJ, Boekelheide K. Dose-dependent effects of Sertoli cell toxicants 2,5-hexanedione, carbendazim, and mono-(2-ethylhexyl) phthalate in adult rat testis. Toxicol Pathol 2007;35:719–727

Moon HJ, Han SY, Shin JH, Kang IH, Kim TS, Hong JH, Kim SH, Fenton SE. Gestational exposure to nonylphenol causes precocious mammary gland development in female rat offspring. J Reprod Dev 2007;53(2):333–344.

Neal MS, Zhu J, Foster WG. Quantification of benzo[a]pyrene and other PAHs in the serum and follicular fluid of smokers versus non-smokers. Reprod Toxicol 2008;25(1):100–106.

Nusier MK, Bataineh HN, Daradka HM. Adverse effects of propranolol on reproductive function in adult male mice. Pak J Biol Sci 2007;10(16):2728–2731.

Petrillo SK, Desmeules P, Truong TQ, Devine PJ. Detection of DNA damage in oocytes of small ovarian follicles following phosphoramide mustard exposures of cultured rodent ovaries in vitro. Toxicol Appl Pharmacol 2011;253(2):94–102.

Richburg JH, Johnson KJ, Schoenfeld HA, Meistrich ML, Dix DJ. Defining the cellular and molecular mechanisms of toxicant action in the testis. Toxicol Lett 2002;135(3):167–183.

Savitz DA, Arbuckle T, Kaczor D, Curtis KM. Male pesticide exposure and pregnancy outcome. Am J Epidemiol 1997;146: 1025–1036.

Schrag SD, Dixon RL. Occupational exposures associated with male reproductive dysfunction. Annu Rev Pharmacol Toxicol 1985;25:567–592.

Shi Z, Zhang H, Liu Y, Xu M, Dai J. Alterations in gene expression and testosterone synthesis in the testes of male rats exposed to perfluorododecanoic acid. Toxicol Sci 2007;98(1):206–215.

Siu ER, Mruk DD, Porto CS, Cheng CY. Cadmium-induced testicular injury. Toxicol Appl Pharmacol 2009;238(3): 240–249.

Smart RC. *Molecular and Biochemical Toxicology*. 4th ed. Hoboken, John Wiley & Sons Inc.; 2008. p 805–830.

Soares SR, Melo MA. Cigarette smoking and reproductive function. Curr Opin Obstet Gynecol 2008;20(3):281–291.

Sobinoff AP, Beckett EL, Jarnicki AG, Sutherland JM, McCluskey A, Hansbro PM, McLaughlin EA. Scrambled and fried: cigarette smoke exposure causes antral follicle destruction and oocyte dysfunction through oxidative stress. Toxicol Appl Pharmacol 2013;271(2):156–167.

Spira A, Multigner L. The effect of industrial and agricultural pollution on human spermatogenesis. Hum Reprod 1998;13(8):2041–2042.

Stouder C, Paoloni-Giacobino A. Transgenerational effects of the endocrine disruptor vinclozolin on the methylation pattern of imprinted genes in the mouse sperm. Reproduction 2010;139(2):373–379.

Turner P. Recent observations on drugs and human fertility. Postgrad Med J 1988;64(754):578–580.

Wong CH, Cheng CY. The blood-testis barrier: its biology, regulation, and physiological role in spermatogenesis. Curr Top Dev Biol 2005;71:263–296.

White SS, Calafat AM, Kuklenyik Z, Villanueva L, Zehr RD, Helfant L, Strynar MJ, Lindstrom AB, Thibodeaux JR, Wood C, Fenton SE. Gestational PFOA exposure of mice is associated with altered mammary gland development in dams and female offspring. Toxicol Sci 2007;96(1):133–144.

White SS, Stanko JP, Kato K, Calafat AM, Hines EP, Fenton SE. Gestational and chronic low-dose PFOA exposures and mammary gland growth and differentiation in three generations of CD-1 mice. Environ Health Perspect 2011;119(8): 1070–1076.

Wyrobek AJ, Watchmaker G, Gordon L, Wong K, Moore D, II, Whorton D. Sperm shape abnormalities in carbaryl-exposed employees. Environ Health Perspect 1981;40:255–265.

Xia Y, Cheng S, Bian Q, Xu L, Collins MD, Chang HC, Song L, Liu J, Wang S, Wang X. Genotoxic effects on spermatozoa of carbaryl-exposed workers. Toxicol Sci 2005;85(1):615–623.

13

DEVELOPMENTAL TOXICOLOGY

VINCENT F. GARRY AND PETER L. TRURAN

There is increasing recognition that early life stages, including fetal development, can be particularly vulnerable to adverse effects from drugs and chemicals. During early life, toxicity from even transient exposures can be profound, long-lasting, and in some cases, trans-generational. Developmental toxicity focuses on the study of birth and developmental anomalies associated with pharmacologic and environmental agents, and the underlying biochemical and molecular mechanisms.

This chapter discusses:

- Principles of teratology and the prevalence of birth defects
- Physiological factors during pregnancy that influence developmental toxicity
- Evidence for developmental effects resulting from exposure of pregnant women in the workplace
- Methods to evaluate teratogenic potential of drugs and chemicals in laboratory animals
- Evaluation of potential neurodevelopmental toxicity
- The role of epigenetic effects in developmental toxicity

13.1 BIRTH AND DEVELOPMENTAL ANOMALIES: TERATOLOGY

Josef Warkany (1902–1992) is described by his peers as the father of modern teratology as a clinical and experimental discipline. His life's work grew from discoveries of the causal role of nutritional and vitamin deficiencies in birth defects, to explorations of the teratologic effects of drugs (warfarin, salicylates, thalidomide, methotrexate, and aminopterin). His ideas were merged with those of James G. Wilson, who formulated the central working paradigm of teratology in the twentieth century, and which is defined as the principles of teratology.

In the period from 1900 through the 1960s, experimental embryology came to the forefront. Developing embryos in sea urchins, amphibians, and chicks were painstakingly dissected, and embryonic tissues at early stages of development were subjected to experimental manipulation. Studies by Weismann, Spemann, and Horstadius provided precious insights into cellular development and organ system differentiation. In parallel with these studies, Horstadius and others explored the effects of chemicals and alterations in the physical environment of the developing embryo.

The Principles of Teratology

The six principles of teratology put forward by James G. Wilson (1915–1987) integrated the knowledge gained from developmental biology, genetics, clinical medicine, and biochemistry to describe the major factors that may be involved in a teratologic event.

These principles, which continue to be relevant today, are as follows:

1. Susceptibility to teratogenesis depends on the genotype of the conceptus and the manner in which this interacts with environmental factors.
2. Susceptibility to teratogens varies with the developmental stage at the time of exposure.

Principles of Toxicology: Environmental and Industrial Applications, Third Edition. Edited by Stephen M. Roberts, Robert C. James and Phillip L. Williams.

3. Teratogenic agents act in specific ways on developing cells and tissues to initiate abnormal developmental processes.
4. The ultimate manifestations of abnormal development are death, malformation, growth retardation, and functional disorder.
5. The access of adverse environmental influences to developing tissues depends on the nature of the agent.
6. Manifestations of deviant development increase in frequency and in degree, as dosage increases, from no effect to the 100% lethal (LD100) level.

Birth Defect Prevalence

More than 3 in every 100 live births will result in a child with a major structural or genetically based birth defect. Birth defects are a major contributor to infant mortality and result in billions of dollars in health care costs. In the United States, national estimates of birth defect prevalence are based on birth certificate data. Based on these data, the overall birth defect rate per year from 1999 to 2005 was approximately 1 in 100 live births. However, birth certificate data underestimate the frequency of birth defects, since not all birth defects are apparent and reported at birth or shortly thereafter. To illustrate the point, the relative rate of birth defects in a cohort of members of the farming community reported in the first year of life, were examined. Comparison of birth defects rates confirmed in medical records versus those from birth certificate data demonstrated that the relative rate per 1000 births of CNS (4.6 vs. 1.3), musculoskeletal (9.8 vs. 6.5), and cardiovascular (5.2 vs. 3.7) birth defects were higher in the medical record confirmed data set. On the other hand, gastrointestinal (2.0 vs. 1.3), urogenital (4.0 vs. 4.4), and genetic/metabolic (0.65 vs. 1.7) birth anomaly rates were nearly equivalent. It is apparent from these data that, although birth anomalies visible at birth or symptomatic shortly thereafter can be captured in birth certificate data, many birth and developmental anomalies make their appearance well after the newborn period. For example, in the same study, 62% of the birth anomalies reported were identified in the first year of life, another 10% within years 1–3, and 20% in children more than 3 years old. More comprehensive work is ongoing through the Centers for Disease Control (CDC).

For more than 30 years metropolitan Atlanta, Georgia, through the CDC, has engaged, and continues to develop, an active birth defect reporting system. Medical records of each live birth in the five counties of metropolitan Atlanta are examined and updated through age 5 by trained abstractors. Major structural and or genetic defects are identified in this group. From 1978 to 2005 the overall prevalence of major structural defects was stable, varying from 2.8 to 3.0 per 100 live births. Male children had more defects than females; fewer birth defects were observed among minorities. Issues such as access to care and poverty appear unresolved in these assessments.

Active reporting of birth defect prevalence nationwide is limited to specific major anatomic disorders. However, the cause(s) of approximately 70% of birth defects reported are unknown, and there are limited data available on the frequencies of genetic, environmental, infectious, and medication-related causes of birth defects. The spectrum of agents known to cause human birth defects is broad. A comprehensive catalogue of known human teratogenic agents was assembled by Shepard in order to capture the available knowledge base regarding agents linked to human birth defects. Nevertheless, based on our current assessments, it is likely that both the spectrum of agents and frequency of birth defects are wholly underestimated, particularly in relation to workplace exposures.

Given the numbers of women employed outside of the home, it is appropriate to make an assessment of pregnancy and reproductive health issues related to the workplace. We will also explore adverse birth and developmental risk.

13.2 PHYSIOLOGY OF PREGNANCY

Pregnancy is a complex physiologic condition that involves dramatic changes in the developing fetus and placenta, and in the cardiovascular, neuroendocrine, renal, respiratory, and metabolic systems of the pregnant woman. All these changes take place in concert as each

TABLE 13.1 Predictive Value of Animal Teratogenesis Studies of Known Human Teratogens ($N = 35$ Schradein 1993)

	No. of Substances	True Positives (%)	Indeterminate (%)	False Negatives (%)
Rat	31	61	10	29
Rabbit	27	41	3	56

Source: Adapted from analysis by Bailey (2005) of Schradein (1993) data.
After analysis, the data provided are categorized as positive, indeterminate, and negative according to the studies reviewed for each of the two species tested. The percent (%) of true positive and false negative are listed. These data suggest that the predictive value of whole-animal teratology studies is limited.

trimester of pregnancy proceeds. In this milieu, xenobiotic chemical and pharmacologic agents, and physical and psychological stressors, have unique opportunities to exert adverse effects through dysregulatory and cytotoxic mechanisms.

The first trimester is the time of structural development of major organ systems of the embryo. The timing of cell movements, and the homing of cells to their developing organ sites, involves exquisitely integrated morphogenic events. For example, formation of the kidney involves bringing together the structural elements of the glomeruli, the parenchyma of the organ, along with the formation of renal tubules and the collection system of the kidney. As pregnancy proceeds, the fetal kidney develops and evolves along phylogenetic lines from the primitive pronephric kidney, through the mesonephric kidney, to the mature metanephric kidney. Accordingly, this period represents a time of unique sensitivity to relatively low-dose chemicals, structural teratogens, whose mechanism of action will interfere with development and differentiation of the organ systems.

Maternal Cardiovascular System

During the first 4–6 weeks of gestation there are dramatic changes in maternal hemodynamics including increased cardiac output, expansion of the plasma volume, and reductions of vascular resistance and arterial pressure. By late pregnancy, the volume of circulating blood has increased by 40–45%. Vascular resistance remains low due to reduced sensitivity to angiotensin and increased nitric oxide levels. The end result is a relative lowering of serum albumin and other serum protein levels. As a consequence, the ratio of protein-bound drugs/toxicants to free drug/chemical levels is altered; more free-drug/toxicant is then available for hepatic biotransformation and/or renal excretion.

Metabolism in the Prospective Mother

The metabolic activities of many of the P450 enzymes are increased during pregnancy. In particular, among the 57 gene products responsible for the expression of the broad range of P450 enzymes, cytochrome P450 3A4 (CYP3A4) is induced to higher activity. This enzyme plays a major role in the metabolism/detoxification of a wide variety of xenobiotics, including approximately 50% of therapeutic drugs. Increased transformation of xenobiotics, toxins, and chemicals with mutagenic potential results in their rapid clearance from the maternal circulation, and hence lessens the amount of toxic agent available for fetal exposure. The metabolic activities of other P450 enzymes (CYP2D6, CYP2A6, CPY2C9) are similarly enhanced. For example, metabolism and clearance of both nicotine and cotinine are enhanced during pregnancy through increased CYP1A2 activity. The metabolic activity of some other P450 enzymes including CYP1A2 (e.g., caffeine metabolism) and CYP2C19 are reduced during pregnancy.

Metabolism in the Fetus

The liver of the developing fetus also contributes to CYP450-mediated metabolism. Among the CYP450 enzymes present in fetal liver, CYP3A7 appears early in development and is thought to play a major role in fetal metabolism of xenobiotics including the metabolic activation of Aflatoxin B1. It is noteworthy that, immediately after birth, CYP3A becomes the predominate isoform in the metabolism of xenobiotics.

Metabolism in the Placenta

During pregnancy, the human placenta is an important endocrine organ. Placental metabolic capacity is mainly directed toward synthesis and metabolism of hormones (both steroidal and protein). Smoking, xenobiotics such as TCDD, and other endocrine disruptors are thought to affect placental steroidogenesis. As the third member of the CYP450 xenobiotic metabolic triumvirate, the trophoblastic placenta is believed to play a fetoprotective role in the first trimester of pregnancy. CYP1A1 is the only placental xenobiotic metabolizing enzyme whose expression and inducibility have been demonstrated throughout pregnancy. Products of cigarette smoking and the pro-carcinogenic PAHs undergo bioactivation by CPYA1. Elevated CYPA1 activity has been associated with adverse birth outcomes such as premature birth, intrauterine growth retardation, and structural abnormalities.

Maternal Respiration

Minute ventilation (FEV1) is increased by the 7–8 weeks pregnancy with an increased tidal volume of 40% and decrease residual volume of 20%. This increased respiratory rate, accompanied by higher cardiovascular output, leads to enhanced pulmonary absorption of air pollutants, including cigarette smoke. More importantly, adaptation of the respiratory system provides enhanced oxygenation for prospective mother and fetus.

Maternal Renal Function

The renal glomerular filtration rate begins to increase in the first trimester of pregnancy and peaks in the second half of pregnancy at a level 40–60% higher than nonpregnant women. An increased excretion of low-molecular-weight water-soluble toxicants can be expected.

Summary of Pregnancy Physiology

The extraordinary changes of pregnancy are designed to assure development and survival of the fetus through physiologic adaptation of the major organ systems and metabolism of the prospective mother. At the same time, these postconception physiologic changes can constitute individual and collective risk factors for teratologic and other adverse reproductive effects caused by chemical and other stressors.

13.3 WORKPLACE, OCCUPATION, AND ADVERSE REPRODUCTIVE EFFECTS

According to the U.S. Department of Labor statistics (2010), about 59% of women residing in the United States work for pay outside the home. Nearly three of every four women who are employed are of reproductive age. According to one source, fully 80% of working women will become pregnant during their working lives. Approximately, 40% of women who were, or became, pregnant while working remained in the workforce. In fact, a woman's right to be pregnant in the workplace is protected by law. As an extension of the Title VII Civil Rights Act (1964), the Pregnancy Discrimination Act of 1978 stipulated that all employers treat pregnant and nonpregnant employees in the same way, both in terms of benefits received and in all other respects. The Family and Medical Leave Act (FMLA) 1993 further stipulated that men and women might take as many as 12 weeks of unpaid leave annually for the birth or adoption of a child. In other countries (European Union, Canada), family leave includes financial and social support.

Given the breadth and scope of this demographic information, it is remarkable that there are so few concrete human data on the risks for birth defects in occupations where women of reproductive age are employed and exposed to toxicants. According to epidemiologic reviews, convincing evidence linking occupational exposure during pregnancy and birth defects is lacking. It is not surprising, therefore, that in the occupational health sector and in the legal arena, strategies and actions devoted to birth defect prevention are limited. In guidance provided by the American College of Occupational and Environmental Medicine indicates, needs assessment points to the application of technologies, including medical surveillance, biomarker exposure/risk assessments, and population-based epidemiology in order to detect persons at risk for birth defects. These are key elements in birth defect preventative strategies but, to date, progress toward their implementation has been notably short of ideal.

Nevertheless, certain maternal occupations can be established as risk factors for birth defects. Data from the national birth defects prevention study show that women employed in any one several occupational groups (janitors/cleaners, scientists, and electronic equipment operators) have an excess risk for birth defects. Other published works provide some detail regarding exposure to chemical agents linked to birth defects.

Nursing/Health Care Workers

Children born to nurses exposed to anesthetic gases may be at increased risk for birth defects, as are those handling cytostatic agents. In general, nurses may have a modest excess risk.

Maternal and Paternal Exposure to Organic Solvents

Solvent exposure in the workplace, mainly for those employed as painters, has been associated with increased risk of birth defects. For males, employment as a painter 3 months prior to spouse's pregnancy was associated with increased risk of birth defects. For females, exposure to solvents in the first trimester of pregnancy has been associated with increased risk of bearing children with cardiac birth defects. Solvents and other products classified as endocrine-disrupting chemicals can be risk factors for male genital malformations.

Radiation

Surprisingly, genetic studies of the offspring of atomic bomb survivors (Hiroshima/Nagasaki, Japan) conducted since 1948 have yielded little or no evidence of excess birth defects. Future molecular genetic studies of the survivors are expected to yield information on transgenerational changes in the genomes of their offspring including changes affecting adult-onset multifactorial diseases, e.g. diabetes. In studies of chronic low-level radiation, such as occurs in the nuclear industry, women radiation oncologists exposed during their pregnancy show little or no difference in the frequency of birth defects. One small study, where maternal occupational exposure to radiation occurred in the first trimester of pregnancy, suggests a significant excess of birth defects in children from radiation-exposed mothers compared to unexposed mothers. In contrast to available data on humans, animal studies precisely document radiation doses, timing during pregnancy, and types of birth defects that may occur.

Finally, rules for radiation dose limits apply once a pregnant employee declares her pregnancy. The code of federal

regulations set an occupational exposure limit of 0.005 Gy (0.5 rad).

Perspective from Studies of Pregnant Workers

Despite the limited demonstration of birth defects causally linked with occupational exposures, the possibility of such effects remains a concern. Clarification of current occupational epidemiologic data, with continued data collection on pregnancy, occupation, and the utilization of the newer pharmaco-genetic approach, will yield concrete evidence-based advice to the community at large.

13.4 GENERAL TESTING METHODS FOR TERATOGENIC EFFECTS IN EXPERIMENTAL ANIMALS

Most human teratogens have been identified by astute clinicians and/or through epidemiologic study. That being said, animal studies historically have been, and are currently, the mainstay of nonclinical investigative drug development and environmental chemical studies of reproductive toxicity. The typical approach to animal chemical teratogenicity assessment is shown in Figure 13.1. In this scheme test chemical is administered to the dam at the time of implantation through the period of organogenesis and organ system development. Necropsy is conducted on the days specified, adjusting for gestation time per species. This is but one part of the developmental toxicologic testing proscribed by U.S. Federal Regulation (Title 40 subpart H 799.9370 TSCA prenatal developmental toxicity) and by OECD guidelines (Guideline 414 prenatal toxicity study). On the whole, both regulatory program assessments are quite similar. They are highly detailed in their study requirements. We will limit our discussion to the basic requirements of the test:

1. Two species should be used, one rodent (rat, mouse) and one nonrodent (rabbit), but commonly the rat is the only species examined.
2. Twenty young adult females per chemically exposed group and concomitant controls are used. Only animals with implantation sites at necropsy can be considered as part of the study group.
3. A minimum of three dose points should be used, these being selected by initial screening to identify dose levels giving approximately 10% maternal toxicity as the highest dose with incremental lowering of each dose by two- to fourfold to approximate the NOEL (no observed adverse effect level).
4. The test substance is usually administered orally by intubation at the same time of day.
5. At necropsy, fetuses are examined for skeletal and soft tissue abnormalities.
6. Data from the individual animal and group are reported in a standard format.

Despite the detailed study protocol briefly recounted here, the overall probability of predicting a human teratogen in standard animal (rat, rabbit) reproductive studies is only moderate (Table 13.1). Increasing the number of species tested (upward of five or more) improves the overall capture rate to almost 100% positive correlation with known human teratogens. While the possibility of animal-based precise

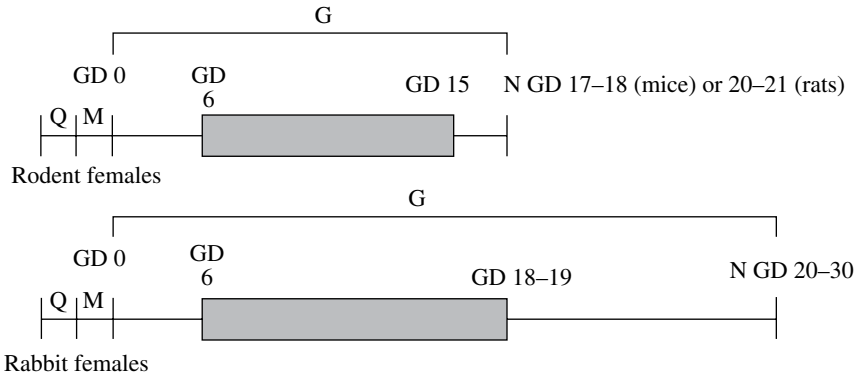


FIGURE 13.1 Scheme for chemical teratogenicity assessment. G, gestation; GD, gestational day; L, lactation; M, mating; N, Necropsy; Q, quarantine; W, wean.

predictive developmental toxicology exists, the overall economic and ethical costs are prohibitive.

To deal with this conundrum requires innovation. A number of approaches have been devised, in large part complementary to each other. These efforts include:

1. Cataloging chemicals by structure–activity relationships to establish chemical structures that are likely or unlikely to be teratogens
2. Utilizing a pharmacokinetic approach to establish a relevant dose mimicking human exposure
3. *In vitro* programming of human stem cells, embryonic cell aggregate formation, and cultured animal embryos
4. Use of nonmammalian species including Zebra fish, frog, sea urchin, and chick embryo
5. Development and use of molecular mechanism based on high-throughput screening profiles of chemicals to be studied paired with computational database-driven toxicology

Integration and execution of all aspects of the strategic effort listed above are part of the European Union REACH (Registration, Evaluation, Authorization and Restriction of Chemicals) program and in the US Tox 21 (toxicology in the twenty-first century), a partnership of the National Toxicology Program (NTP), National Institutes of Health Chemical Genomics Center (NCGC), and the U.S. EPA. The vision and concepts embodied in these efforts are designed to utilize crossover data from multiple disciplines to minimize animal usage and maximize more insightful and cost-efficient assessment of chemicals in current use. Research into toxicologic properties of pharmaceuticals and industrial products in the planning and early stages of R&D are expected to be optimized. Included in this paradigm is an evolving understanding of embryonic growth and development.

In sum, chemicals that cause perturbation of defined mechanisms of normal growth and development are expected to be recognized as potential human teratogens.

13.5 NEURODEVELOPMENTAL TOXICITY

The developing nervous system can be particularly sensitive to the effects of drugs and chemicals with long-lasting consequences. As a result, there has been considerable interest in teratology related to neurobehavioral effects. The history and purview of neurobehavioral teratology is intertwined with neurodevelopmental toxicology. The broad field of behavioral teratology includes behavioral changes produced by toxicants with documented anatomic neuropathologic findings; for example, methylmercury, ethanol, parathion, and others where there is significant CNS change, but little if any change in the behavior of the offspring of treated animals. In other instances, behavioral changes may be delayed well into the postnatal period and beyond through adulthood, without marked neuroanatomic findings, for example in utero lead and PCBs.

As early as 1963, the concept of behavioral teratology was introduced in a review by Werboff and Gottlieb. In that review, the authors described the postnatal behavioral effects of exposure to prenatal X-radiation and to the use of psychoactive drugs in pregnancy. After a period of more than 10 years, Great Britain and Japan incorporated guidelines for neurodevelopmental toxicity testing (DNT) of medicinal products. The 1980s saw the development of the first draft of the U.S. EPA DNT protocol, and over time, behavioral testing procedures have evolved as a cooperative effort among agencies and scientists representing the United States, Japan, and European Union.

The current U.S. EPA neurodevelopmental toxicity study guidelines for studies in animals are shown in Figure 13.2. The figure outlines parameters of observation and times for specific neurobehavioral examinations for a typical DNT study. The time frame examined extends from day 6 of gestation to the 60th day after birth. Dosing of animals starts at gestational day (GD) 6 and continues through lactation (postnatal day 21; PND21) in the developing/maturing rat. Study parameters are comparable to OECD test guideline 426 (adopted October 2007). The dose response curves produced using the behavioral endpoints specified tend to be complex, and may involve a U-shaped dose response (e.g. alcohol increases motor activity at low doses with CNS depression, but induces decreased motor activity at high doses). It is of interest that no observable effect (NOEL) levels observed in these behavioral studies tend to occur at lower levels than in routine teratologic studies. Conducting these animal studies in a regulatory context is expensive and time-consuming, but survey of their use and value in safety assessment studies suggests that behavioral parameters can be an improvement over more general toxicologic assessments.

In vivo and *in vitro* screening of potential neurotoxicants, utilizing appropriate biomarkers and toxicant pathway analysis, offer alternatives which allow us to step to the future and address the onslaught of untested chemicals with the potential for harm. While these efforts contribute to our understanding of neurodevelopmental toxicants, the historic insights gained from clinical observation coupled with behavioral biomarker studies should not be overlooked.

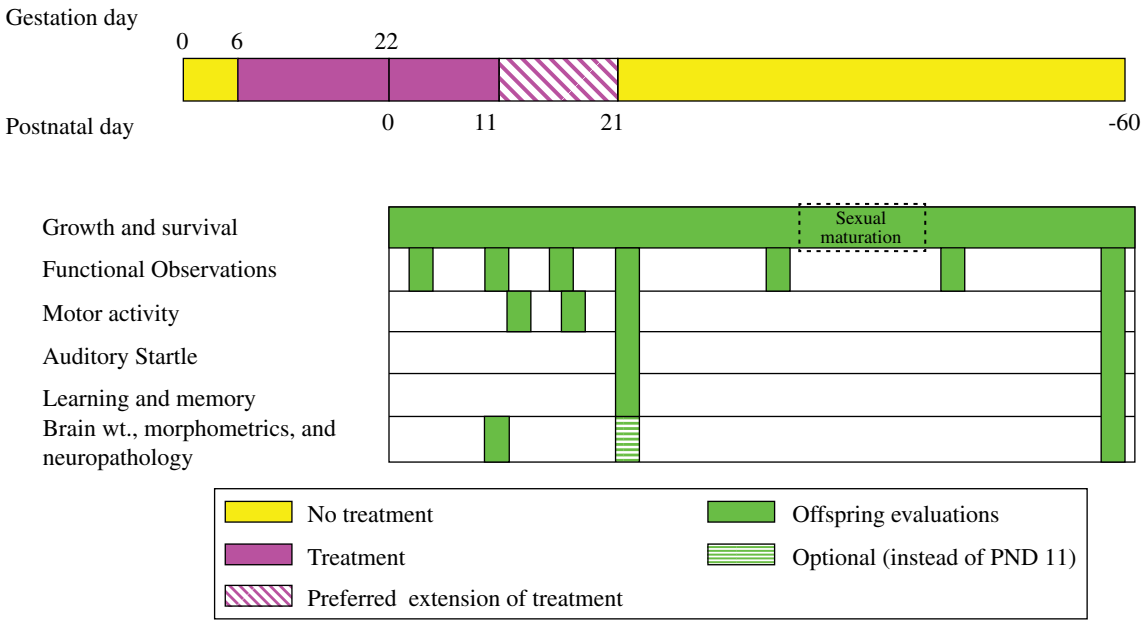


FIGURE 13.2 Study design for developmental neurotoxicity testing. *Source:* From Raffaele et al. (2010), with permission. © 2010 Elsevier.

13.6 THE ROLE OF EPIGENETICS IN ENVIRONMENTAL EFFECTS

Brief History of Epigenetics

In the 1940s and 1950s, Waddington attempted to join the findings of developmental biology with those of genetics. He coined the term "epigenetics" derived from joining the term genetics with "epigenesist." This latter term was used to describe the differentiation of cells from their totipotent state in embryonic development. In the 1960s, converging along these same lines in the field of cancer biology, GB Pierce used the teratocarcinoma cell model to establish that single undifferentiated cancer cells grown *in vitro* can differentiate to become nerve cells, bone, muscle, or connective tissue. He postulated that the cancer cell is a pluripotent "stem" cell. Furthermore, terminally differentiated cells can also be reprogrammed into stem cells or to other cell types, for example skin fibroblasts to cardiomyocytes, apparently without a change in DNA sequence.

As molecular genetics advanced, more refined definitions of the phenomenon of programming and reprogramming differentiation were proposed. In 1990, Holiday defined epigenetics as "the study of the mechanisms of temporal and spatial control of gene activity during the development of complex organisms." As the programming process was further studied, and its multigenerational implications clarified, the current definition of epigenetics was established as including molecular factors and processes around DNA that are mitotically stable and regulate genome activity independent of DNA sequence.

Epigenetic Mechanisms

The most commonly studied molecular epigenetic mechanisms include changes in DNA methylation, histone modification, chromatin remodeling, and posttranscriptional alteration of gene expression mediated by noncoding RNAs. Of these mechanisms, DNA methylation and histone modification are best understood. We will focus our attention on these areas and their developmental toxicological impact.

DNA Methylation In the mammalian genome, DNA methylation occurs predominately by covalent modification of the fifth carbon of the cytosine base. The majority of these modifications are present in the CpG regions of DNA where a cytosine nucleotide occurs next to a guanine nucleotide in the linear sequence of bases. The methylated form of the DNA base cytosine (5-methylcytosine) represents about 1% of the total DNA bases and is estimated to represent 70–80% of all the CpG dinucleotides in the genome. They are scattered throughout the genome, concentrated in pockets (CpG islands). Mammalian DNA methylation has a crucial role in maintaining pluripotency, X chromosome inactivation, and genomic imprinting. It also protects genomic integrity by silencing transposable elements to insure chromosome stability. In routine cell division, the enzyme DNA methyltransferase 1 performs a maintenance role by propagating DNA methylation marks from the old strand to the new DNA strand. In this sense, DNA methylation marks are known to be heritable.

Epigenetic programming of the germ line occurs during the migration of the primordial germ cells in the embryo. The migrating primordial germ cells in the genital ridge undergo an erasure of methylation (global demethylation) and colonize the early bipotential gonad before gonadal sex determination. Once sex determination is initiated, the primordial germ cells develop female or male germ cell lineage and remethylate the DNA in a male- or female-specific manner. De novo methylation is mediated by DMNT3 enzymes. During the period of sex determination, the epigenome is most sensitive to toxicants that may lead to transgenerational toxicant effects. Recently, rare regulatory elements have been identified that escape DNA demethylation providing a potential mechanistic basis for transgenerational inheritance.

Histone Modification and Chromatin Function Chromatin is the packaging material for DNA. It is organized into two general structures, heterochromatin and euchromatin. Heterochromatin is generally silent, and is located at centromeres and telomeres of chromosomes. It tends to be rich in repetitive DNA sequences. Euchromatin contains most of the transcriptionally active DNA. Nucleosomes are the basic unit of chromatin consisting of 147 bp of DNA wrapped around a histone octomer.

Histones may be modified by processes that include acetylation, methylation, ubiquitination, sumoylation. The most common modifications are acetylation and methylation of the lysine residues in the amino terminal of histone 3 and histone 4. Increased acetylation induces transcriptional activation; decreased acetylation induces transcriptional repression. Methylation of histones may be associated with repression or activation depending on the location of the lysine residue. Apart from alteration gene expression or repression, histone modification appears to play an important role in other cellular processes such as response to DNA damage. Dysregulation of histone modification of the male gamete may play a pivotal role in male-mediated developmental toxicity induced by cyclophosphamide.

Evidence for Epigenetic Effects on Lifelong and Multigenerational Phenotypic Change

Nutrition and Diet World War II provided a natural experiment of the relationships between diet and multigenerational health outcomes. Survivors of the Dutch famine (1944–1945) are one well-studied example. During this time frame, registries and health care in the Netherlands remained intact, and official food rations were documented. As a result, the population could be epidemiologically assessed. Prenatal exposure to famine was associated with increased metabolic disease and cardiovascular disease in the offspring. Offspring (grandchildren) of prenatally undernourished fathers, but not mothers, were heavier and more obese. These persistent health outcome differences indicated a possible epigenetic effect. IGF II (insulin-like growth factor 2) is a key factor in growth and development. It is maternally imprinted and is differentially methylated. The offspring of individuals prenatally exposed to the Dutch hunger winter had less DNA methylation than their unexposed same-sexed siblings six decades later. In a somewhat similar study of 60 exposed siblings with unexposed same-sexed siblings, prenatal famine and genetic variation showed a similar association with IGF2/H19 methylation. Their contributions were additive. Methylation abnormalities at the *IGF2-H19* gene are associated with congenital growth disorders and Wilm's tumor.

In a detailed study of methylation patterns at birth, microarray technology was used to direct gene site selection for examination of the possible relationship with childhood obesity in two independent cohorts. Increased methylation of the retinoid X receptor-alpha (RXRA) was correlated with a lower maternal carbohydrate intake early in pregnancy, higher neonatal adiposity, and childhood obesity (children aged 6 and older). RXRA is involved in insulin sensitivity, adipogenesis, and fat metabolism. DNA methylation profiling identified epigenetic dysregulation in the pancreatic islets from type 2 diabetics, but not from peripheral blood, indicating some tissue-specificity for epigenetic effects.

There are limited data regarding the role of nutrient-induced methylation. It appears that dietary supplementation or deficiencies of nutrients can result in hypermethylation of certain gene sites while not in others. Studies of folate use after the 12th week of pregnancy, and the imprinted *IGF2* gene, showed increased methylation while the LINE-1 (long-dispersed nuclear element) was decreased. The consequences of these epigenetic nutritional effects are not understood. Interestingly, studies of Bisphenol A (BPA) exposure in a mouse model show that BPA induces hypomethylation. This effect was counteracted by dietary supplementation with methyl donors (B_{12}, folate, and choline) or genistein. In the Agouti mouse model, dietary supplementation with methyl donor nutrients increased methylation and darkened coat color. Coat darkening persisted into the next generation.

The concept of environmental factors, acting early in life, leading to predisposition to disease later in life was given an epidemiologic basis by the works of Barker and coworkers. In these seminal works, low-birth-weight babies who survived to adulthood were found to be at special risk for coronary artery disease, hypertension, and type 2 diabetes. These findings, and the current epidemic of obesity, eventually led to the establishment of the OBELIX project (OBesogenic Endocrine disrupting: Linking prenatal eXposure to the development of obesity later in life). This is a 4-year project begun in 2009 and designed to assess early-life exposures to endocrine-disrupting chemicals (EDCs) using mother–child

cohorts from four European regions (N ~6000). Clinical biomarkers, concomitant animal studies, and prenatal integrated food-contaminant/dietary assessment will form the basis for risk assessment of potential obesogenic EDCs.

Some obesogenic endocrine disruptors also induce epigenetic transgenerational inheritance of obesity and reproductive disease through sperm DNA methylation epimutations. These environmental toxicants include Bisphenol A, Bis (2-ethylhexyl) phthalate (DEHP), and dibutyl phthalate (DBP). Other transgenerational effects have been reported for the fungicide Vinclozolin, including testicular, prostate, and kidney abnormalities, and polycystic ovaries. Somewhat similarly, dioxin-induced total disease and multiple disease increased from the F1 to the F3 generation. Prostate disease and kidney disease in males, and pubertal abnormalities and polycystic ovaries in females, were among the panopoly of diseases identified in excess. Fifty differentially methylated regions were identified in gene promoters in the sperm epigenome.

In a different perspective, prenatal exposure to the fungicide tributyltin (TBT) was found to modulate critical steps of adipogenesis through activation of the peroxisome proliferator gamma receptor. Mesenchymal stromal cells were reprogrammed to adipocytes *in vitro* and *in vivo*. From these initial studies was presented the hypothesis that exposure of the epigenome of multipotent mesenchymal stem cells to obesogens induced developmental bias in favor of adipocyte transformation at the expense of bone formation. More recently, prenatal TBT exposure was found to produce transgenerational effects on fat depots. Increased fat depot size, mesenchymal cell reprogramming, and hepatic steatosis were noted. Curiously, liver morphology described in these studies was consistent with nonalcoholic fatty liver. Human epidemiological studies have variously associated obesity and/or diabetes with environmental chemical exposure, including arsenic, TCDD, DDE, POPs, PCBs, organochlorine pesticides, nonylphenol, PFOA, and BPA.

Cardiovascular Disease and Hypertension Cardiovascular disease and disease pathways are at the leading edge of epigenetic investigations. Within this approach lies the possibility of new pharmacologic, dietary, and behavioral therapies for cardiovascular disease and type 2 diabetes. Several epigenetic biomarkers are emerging: LINE 1 (Interspersed repetitive elements) methylation status and the S-adenosylmethionine (SAM)/(SAH) S-adenosylhomocyteine ratio. SAM and its conversion to SAH are indicators of methyl perturbations. As such they may serve as indicators of exposure and predictors of cardiovascular risk. In a study of ageing adults exposed to carbon black (particulates), increasing PM 2.5 levels were associated with decreased LINE-1 methylation. Lower LINE-1 levels correlated with higher blood pressure. Similarly, LINE 1 was inversely correlated with existing diagnosis of hypertension, ischemic heart disease, and stroke. In occupational studies of welders/boiler makers, methylation at the inducible nitric oxide synthase gene site (involved in the production of nitric oxide important in cardiovascular function, e.g., vasodilation) demonstrated an association with increased methylation in the promoter region of iNOS and with number of years worked in the trade.

Prenatal exposure to DDT in the 50s and 60s was noted to lead to increased prevalence of hypertension among women arising from these births before age 50, suggesting an association with early development.

During the postnatal period, Histone deacetylase(s) (HDACs) have been reported to be key regulators of angiogenesis. Ischemia is a potent inducer of HDAC activity, an example of epigenetic dysregulation, and is a marker of the extent of a myocardial infarct. Inhibition of HDAC by the pharmacologic agent trichostatin (TSA) preserves myocardial performance and prevents remodeling through stimulation of endogenous angiogenesis. These data point to the use of epigenetic mechanisms as part of a therapeutic approach to the treatment of cardiovascular disease.

Neurodevelopment and Stress The hypothalamic-pituitary-adrenal axis (HPA) is a critical feature to our response to physical or social stressors. Dysregulation of HPA function is a common underlying feature in neurodevelopmental and psychiatric disorders. Hypo- and or hyperreactivity of the hormones of the HPA axis have been reported in schizophrenia, autism, and depression. The HPA axis serves as a modulating center governed by positive and negative feedback. Release of stress hormones such as glucocorticoids is regulated at specific brain sites including the hypothalamus, the hippocampus, and the pituitary gland. In mammals, severe chronic stressors during critical periods of gestation have been shown to evoke elevated and prolonged glucocorticoid response in their offspring. Male offspring exposed to stress during early gestation displayed depressive "like" behavior and had elevated stress response as measured by glucocorticoid levels after restraint stress, reduced hippocampal and amydgala DNA methylation at the corticotrophin-releasing gene with elevated corticotropin-releasing hormone (CRH) expression. Sex-specific programming early in pregnancy is thought to contribute to these effects. More recent work has utilized a mouse model in which the investigators have identified a period in early gestation when stress results in dysmasculinized and stress-sensitive male offspring demonstrated transmission of the stress-sensitive phenotype to the second-generation males. miRNA from the brains of these males showed a more female-like pattern. Furthermore, paternal stress alters sperm microRNA content and reprograms offspring HPA axis regulation.

In studies on animal models, subjects prenatally exposed to high levels of maternal stress were obese and showed impaired glucose tolerance as adults. In humans, maternal stressors such as daily social interactions, trauma, or

exposure to natural disasters are associated with elevated basal cortisol levels as well as emotional or behavioral problems. Reduced cognition and risk of diseases such as autism and schizophrenia are increased in their children. The level of maternal care in animals alters glucorticoid receptor gene expression through effects on DNA methylation. Similarly, DNA methylation of glutamic acid carboxylase is affected by level of maternal care in the neonatal period indicating a developmental effect on the GABA (gamma-aminobutyric acid) system.

Substances of Abuse Several animal studies have associated prenatal alcohol exposure with disturbed DNA methylation patterns in the offspring. Maternal ethanol consumption in agouti mice results in CpG methylation at the Avy locus in the offspring, resulting in a shift in coat color, growth restriction, and development of craniofacial deformities. The mechanism(s) for prenatal alcohol effects on methylation are poorly understood and may relate to disturbance on one-carbon (folate, B_{12}, methionine) metabolism.

Chronic administration of cocaine to pregnant mice during early gestation was found to produce DNA methylation changes in the hippocampus of male offspring. A heritable phenotype of cocaine self-administration developed in rats showed delayed acquisition and maintenance of self-administration in male but not female offspring. Brain-derived neurotrophic factor (BDNF) was increased in the male progeny, as was increased acetylation of histone A3. Administration of a BDNF antagonist reversed the cocaine resistance in male progeny. Once again, the epigenetic paradigm may be useful in the development of pharmacotherapy, in this case for cocaine abuse.

The mesolimbic *D2R* gene (dopamine receptor) is vulnerable to cannabis. Specifically, reduction of *D2r* gene expression in the nucleus accumbens (NAc) is affected in the human fetus. Animal studies indicate that reduced D2R expression in the NAc region continues into adulthood. These data, coupled with reduction of histone H3 in the adult animal, suggest that an epigenetic mechanism is involved. The mesolimbic structures (NAc and amygdale) are brain regions important for emotional regulation. In regard to neurite proliferation in the developing animal brain, in cannaboid-exposed animals the neural adhesion molecule L1 mRNA is overexpressed in brain areas where cannaboid receptors are localized. It appears predominantly in male fetuses but not in female fetuses, suggesting that response to cannabis may be dimorphic in some respects. In sum, there is little current indication that cannabis has epigenetic effects.

Cancer Cancer, historically, was thought to be a genetic disease. We now know that epigenetic abnormalities alongside of genetic alteration are partners in the process of carcinogenesis. Abnormal DNA methylation, histone code modifications, and miRNA changes in human cancer cells have all been demonstrated in human cancer cells. It is not possible here to conduct a comprehensive review of the many aspects of the role(s) of epigenetics in cancer. Instead, we will use breast cancer to model various aspects of the epigenetic process in carcinogenesis.

Diethylstilbestrol (DES) and Bisphenol A (BPA) are estrogen-like endocrine disruption chemicals that induce persistent epigenetic changes in the uterus. DES exposure in utero is associated with excess risk of breast cancer (RR = 2.5) in adult women after age 40. The histone methyltransferase histone-lysine *N*-methyltransferase (EZH2) has been linked to breast cancer risk. Overexpression of EZH2 impairs DNA repair and is associated with breast tumor progression and poor prognosis. Mice exposed to DES in utero showed a greater than twofold increase in EZH2 expression in mammary tissue. EZH2 protein was elevated in mice exposed to DES or BPA. Similarly, BPA or DES exposure showed increased mammary histone trimethylation. Increased histone H3 (trimethyl K27) activity at lysine 27 mediates the epigenetic effect of EZH2. In sum, environmental exposure to pseudoestrogens in utero can enhance breast cancer risk through epigenetic mechanisms.

The breast cancer susceptibility gene (*BRCA1*), located on chromosome 17 at Q21, functions normally in DNA repair of double-strand breaks. It is a tumor suppressor gene that fits the two-hit model. According to the model, mutation at one of two alleles confers susceptibility and occurs in the germ cell line. Germ line mutations of *BRCA1* predispose women to breast cancer with a lifetime risk up to 85% by age 70. BRCA mutations vary by ethnic group, both in number and type. The second hit (mutation) occurs in somatic cells in the target tissue and leads to cancer. Although inherited *BRCA1* mutations account for a small portion of breast cancers (2.5–5.0%), a significant portion (30–40%) of the larger group of sporadic (nonhereditary) breast cancers exhibit absence or significantly lower levels of BRCA protein, suggesting epigenetic silencing (epimutation). Loss of BRCA activity inhibits embryonic stem cell differentiation and enhances an aggressive form of breast cancer. Inhibition of *BRCA1* activity (silencing) has been noted after *in vitro* exposure of MCF-7 cells (breast cancer cell line) to the PAH prototype benzo[a]pyrene.

Monozygotic Twins Discordant for Constitutive BRCA1 *Promoter Methylation: A Case Study*

One member of a pair of twin sisters suffered childhood leukemia and a secondary thyroid cancer as an adult. The affected twin displayed increased BRCA methylation (12%) in primary skin fibroblasts; 13% of the BRCA alleles were fully methylated whereas in the unaffected sister there were only single CpG errors. The BRCA epimutation is thought to have originated in an early somatic event carrying one epigenetically inactivated (silenced)

allele. This epimutation was associated with reduced basal *BRCA1* protein level. Other findings included reduced expression (deletion) of the *RSPO3* gene in the affected twin noted in genome-wide microarray. Altogether these data point out a role for epigenetics in *BRCA1* modification of breast cancer risk.

Hypermethylation of BRCA1 *at CPG Islands: A Predictor of Sensitivity to Cisplatin Chemotherapy*

BRCA1 can undergo epigenetic inactivation in sporadic breast cancer through hypermethylation of its promoter-associated CpG island. *In vitro* studies of several breast cancer cell lines demonstrated that those that contained the hypermethylated promoter site were more sensitive to Cisplatin. This potentially extends the utility of Cisplatin beyond its use in hereditary forms of breast cancer to its use in sporadic forms of breast cancer.

It is hoped that the preceding paragraphs point out the diverse role(s) of epigenetics in the identification of the pathologies and treatment approaches for cancer.

13.7 SUMMARY

Over the last 50 years, the science of teratology has emerged from morphologic description of abnormalities of growth and development to a more mechanistic approach. Clinical observation still remains a key element in these assessments. It is the ultimate safeguard against the failure of preclinical studies particularly animal studies to detect the human effects of developmental toxicants such as thalidomide.

SUGGESTED READING

ACOEM Task Force on Reproductive Toxicology. Reproductive and developmental hazard management guidance. J Occup Environ Med 2011;53(8):941–949.

Aguilera O, Fernández AF, Muñoz A, Fraga MF. Epigenetics and environment: a complex relationship. J Appl Physiol 2010;109(1): 243–251.

Bailey J, Knight A, Balcomb J. The future of teratology research is *in vitro.* Biogenic Amines 2005;19(2):97–145

Bale TL, Baram TZ, Brown AS, Goldstein JM, Insel TR, McCarthy MM, Nemeroff CB, Reyes TM, Simerly RB, Susser ES, Nestler EJ. Early life programming and neurodevelopmental disorders. Biol Psychiatry 2010;68(4):314–319.

Bernal AJ, Jirtle RL. Epigenomic disruption: the effects of early developmental exposures. Birth Defects Res A Clin Mol Teratol 2010;88(10):938–944.

Bollati V, Baccarelli A. Environmental epigenetics. Heredity 2010;105:105–112.

Dawson MA, Kouzarides T. Cancer epigenetics: from mechanism to therapy. Cell 2012;150(1):12–27.

Eichner M. Parenting and the workplace: the construction of parenting protections in United States law. Int Breastfeed J 2008;3:14–16.

Garry V, Harkins M, Erickson L, Long-Simpson L, Holland S. Birth defects, season of conception, and sex of children born to pesticide applicators living in the Red River Valley of Minnesota, USA. Environ Health Perspect 2002;11(Suppl 3):441–449.

Garry VF, Truran P. Teratogenicity. In: Gupta RC, editor. *Reproductive and Developmental Toxicology*. London: Elsevier Inc.; 2011.

Granger JP. Maternal and fetal adaptations during pregnancy: lessons in regulatory and integrative physiology. Am J Physiol Regul Integr Comp Physiol 2002;283(6):R1289–R1292.

Hakkola J, Pelkonen O, Pasanen M, Raunio H. Xenobiotic-metabolizing cytochrome P450 enzymes in the human feto-placental unit: role in intrauterine toxicity. Crit Rev Toxicol 1998;28(1):35–72.

Herdt-Losavio ML, Lin S, Chapman BR, Hooiveld M, Olshan A, Liu X, DePersis RD, Zhu J, Druschel CM. Maternal occupation and the risk of birth defects: an overview from the National Birth Defects Prevention Study. Occup Environ Med 2010;67(1): 58–66.

Kim JK, Samaranayake M, Pradhan S. Epigenetic mechanisms in mammals. Cell Mol Life Sci 2009;66(4):596–612.

Kimmel CA. Overview of teratology. Curr Protocols Toxicol 2001;6.13.1:13.1.1.–13.1.8.

Lin S, Herdt-Losavio ML, Chapman BR, Munsiea J-P, Olshanc AF, Druschel CM. Maternal occupation and the risk of major birth defects: a follow-up analysis from the National Birth Defects Prevention Study. Int J Hyg Environ Health 2013;216(3): 317–323. Available at 10.1016/j.ijheh.2012.05.006. Accessed August 4, 2014.

Manikkam M, Tracey R, Guerrero-Bosagna C, Skinner MK. Dioxin (TCDD) induces epigenetic transgenerational inheritance of adult onset disease and sperm epimutations. PLoS ONE 2012;7(9):e46249. doi:10.1371/journal.pone.0046249.

Manikkam M, Tracey R, Guerrero-Bosagna C, Skinner MK. Plastics derived endocrine disruptors (BPA, DEHP and DBP) induce epigenetic transgenerational inheritance of obesity, reproductive disease and sperm epimutations. PLoS ONE 2013; 8(1):e55387. doi:10.1371/journal.pone.0055387.

McElhatton PR. Occupational exposures during pregnancy. J Toxicol Clin Toxicol 2003;41(4):407–409.

Myllynen P, Pasanen M, Pelkonen O. Human placenta: a human organ for developmental toxicology research and biomonitoring. Placenta 2005;26(5):361–371.

Raffaele KC, Rowland J, May B, Makris SL, Schumacher K, Scarano LJ. The use of developmental neurotoxicity data in pesticide risk assessments. Neurotox Teratol 2010;32:563–572.

Salihu HM, Myers J, August EM. Pregnancy in the workplace. Occup Med (Lond) 2012;62(2):88–97.

Schardein JL. Principles of teratogenesis applicable to drug and chemical exposure. In: Schardein JL, editor. *Chemically Induced Birth Defects*. 2nd ed. New York: Dekker, 1993. p 1–63.

Shepard T, Lemire R. *Catalog of Teratogenic Agents*. 11th ed. Baltimore: The Johns Hopkins University Press; 2004.

Spemann H. *ExperimentelleBeitrağezu einer Theorie der Entwicklung. Berlin: Springer. (German edition of the Silliman lectures of 1933): Embryonic Development and Induction. Yale University Press. 1938.* Reprinted: Hafner Publ. Co. New York, 1962.

Tice RR, Austin CP, Kavlock RJ, Bucher JR. Improving the human hazard characterization of chemicals: a Tox21 update. Environ Health Perspect 2013;121(7):756–765.

Warkany J. History of teratology. In: Wilson J, Clark F, editors. *Handbook of Teratology*. Volume 1. New York: Springer-Verlag; 1977.

Weismann A. *Essays upon Heredity and Kindred Biological Problems*. Volume 1. Oxford: Clarendon Press; 1891.

Wilson JG Clarke F, editors. *Handbook of Teratology*. Volume 1. New York: Springer-Verlag; 1977. p 49–62.

14

MUTAGENESIS AND GENETIC TOXICOLOGY

MARTHA M. MOORE, MEAGAN B. MYERS, AND ROBERT H. HEFLICH

Genetic toxicology is the subfield of toxicology that deals with damage to the genetic material. It includes the study of mutations, alterations in chromosome number, damage to chromosomes and primary DNA damage that may or may not ultimately lead to a heritable mutation. Genetic toxicology is a part of applied toxicology. Genetic toxicology data are used in product development, particularly for various hazard assessments, and are either required for, or often available for, use in a number of different types of regulatory decisions. In general, these regulatory decisions can be divided into hazard identification and mode-of-action assessment for chemicals inducing either cancer or mutation. A large variety of products (e.g., pharmaceuticals, pesticides, new use chemicals, food additives, and food contact substances) are tested for their potential to induce genetic damage. The two major human health outcomes of concern are cancer and impaired reproduction (Figure 14.1).

Gene mutations and chromosomal mutations (both structural and numerical), along with cellular toxicities, epigenetic effects, and alterations of gene expression, are all involved in the development of cancer. Germ cell and related tissue mutations can lead to decreased fertility, increased spontaneous abortions and other birth defects, or to new heritable mutations that are passed on to future generations. While it has not been shown that human germ cell mutations can be induced by chemicals and passed on to future generations, it is clear that these effects occur in laboratory animals and it is also clear that new mutations have occurred in human families.

This chapter will cover several aspects of genetic toxicology with a focus on its regulatory use and will be organized as follows:

- Fundamentals of chemically induced genetic damage
- Genetic damage and its impact on human disease
- Genetic toxicology tests
- Regulatory use of genetic toxicology data
- Promising assays under development

14.1 FUNDAMENTALS OF CHEMICALLY INDUCED GENETIC DAMAGE

Deoxyribonucleic acid (DNA) is the fundamental structure that contains all of the information required for cellular function and the information encoded in DNA is passed from one cell generation to the next. DNA is comprised of two antiparallel strands of nucleotides that are linked together to form a double helix. These two strands are held together by hydrogen bonds that form between the complementary nucleotide bases (Figure 14.2). There are four nucleotide bases in DNA (Figure 14.3), two purines (guanine (G) and adenine (A)) and two pyrimidines (cytosine (C) and thymine (T)). To form the double helix, a purine pairs with a pyrimidine. Specifically, A pairs with T and G pairs with C. The ordering of the DNA bases (i.e., the DNA sequence) ultimately controls the function of the molecules that make up cells. The DNA is transcribed into RNA (ribonucleic acid) where uracil is used instead of thymine. RNA serves as the template upon which a series of amino acids are assembled to make a protein. The sequence of amino acids is determined by the order of the nucleotide bases that are decoded in groups of three bases called codons. Each codon indicates a specific amino acid and the order of the inclusion of amino acids into the protein.

When a cell divides, the two strands of DNA separate from each other and, using one of the DNA strands as a template,

Principles of Toxicology: Environmental and Industrial Applications, Third Edition. Edited by Stephen M. Roberts, Robert C. James and Phillip L. Williams.

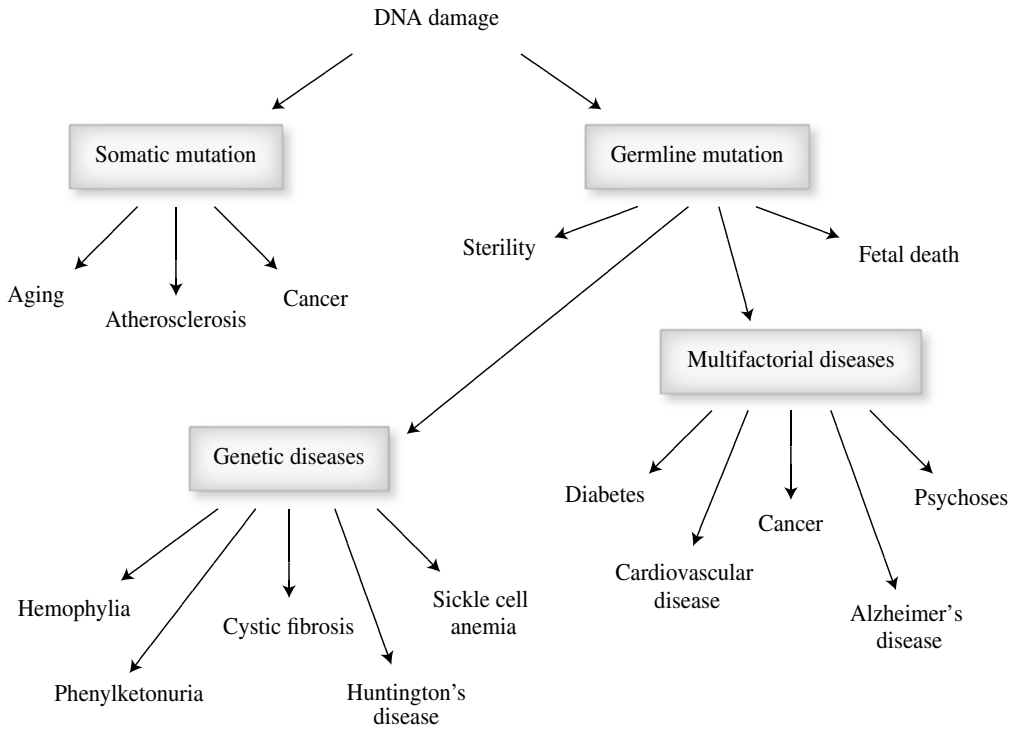


FIGURE 14.1 Mutations can occur in somatic cells or in germ cells and cause a wide variety of adverse health outcomes.

the cellular machinery replicates the strand. When this occurs without error, the daughter cells contain the same DNA sequence as their parent cell. Errors can occur in this replication process, however, and the order of bases can be modified, thus resulting in a mutation (an alteration in the DNA nucleotide sequence order). These types of mutations are "spontaneous" mutations, that is, they occur as a result of the normal process of cell replication and division. Mutations also can occur following exposure to chemicals or physical agents such as radiation.

The DNA is packaged along with proteins into an organized structure called the chromosome. Each chromosome contains a single piece of DNA that includes a large number of genes, regulatory elements, and other nontranscribed nucleotide sequences. Chromosomes are microscopically visible structures and each species has a defined number. Because chromosomes occur in pairs, each cell will contain two copies of each gene, with the exception of the sex determining chromosomes where, for humans (and other mammals), women have two X chromosomes but men have one X and one Y chromosome. These gene copies may not be exactly identical. There can be small variations in the DNA sequence of an individual gene that are not lethal to the cell or organism but that may cause alterations in cellular function. There are a number of human diseases (discussed below) that result from small alterations in genes. These slightly different genes are called alleles. If a particular gene has two identical copies on the two chromosomes, this is designated as homozygous. If the two alleles are different, the cell is heterozygous.

Mutations can be classified as point mutations, gene mutations, or chromosomal mutations; the primary distinction between the various types relates to the number of base pairs altered. Point mutations impact a single base pair and can be base pair substitutions or frame shifts (Figure 14.4). If a purine is exchanged for a purine or a pyrimidine for another pyrimidine, it is designated as a transition. On the other hand, if a pyrimidine replaces a purine or vice versa, it is called a transversion. Base pair substitutions that occur in the coding regions (the part of DNA responsible for ultimately producing a new protein), can be further classified based on their effect on the protein product. A silent mutation occurs when the base pair change does not impact the amino acid order in the protein. This can occur because more than one codon can code for a particular amino acid. A missense mutation occurs when the altered codon designates an incorrect amino acid into the protein and a nonsense mutation codes for a stop codon. In the latter case, the protein will be truncated and will generally not be functional. The three base pair codons consisting of UAA, UGA, or UAG are stop codons, which indicates to the cellular machinery that the protein is complete and no further amino acids should be added. Frame shift mutations involve the insertion or deletion of a small number of base pairs, which impacts the reading frame. That is, it will dramatically affect the three base or codon order, generating a protein containing an

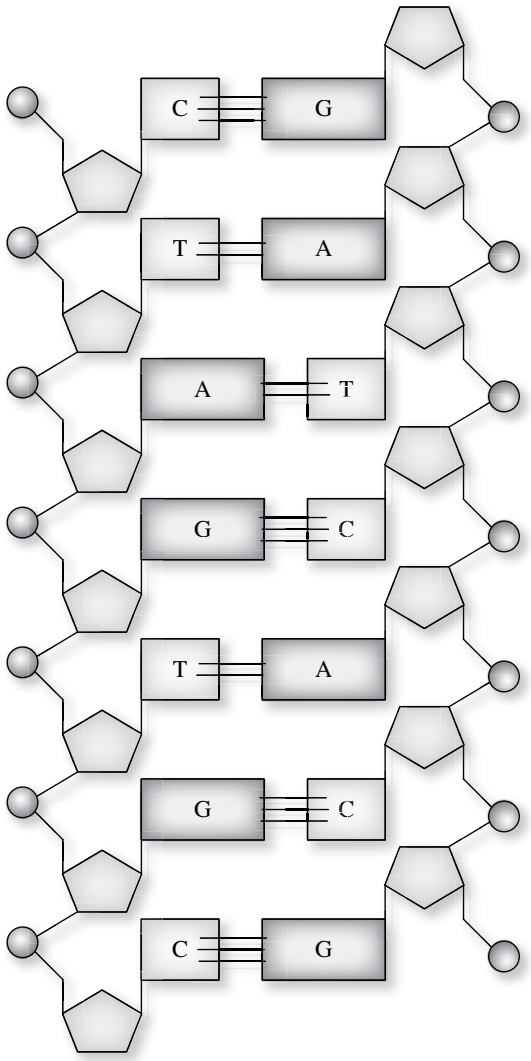


FIGURE 14.2 The DNA double helix is comprised of four different nucleotides that form base pairs held together by hydrogen bonds.

abnormal order of amino acids, which generally results in a premature stop codon at some point in the sequence.

Gene mutations or chromosomal mutations can contain intragenic insertions or deletions. The size of the insertion or deletion will determine whether the impact is to a single gene, a series of linked genes, or a whole chromosome. An intragenic deletion may cause the loss of several introns and exons from a single gene. These deletions may lead to heterozygosity for homozygous alleles or to loss of heterozygosity (LOH) which is the loss of one allele of a heterozygous pair of alleles. A very large deletion can result in the loss of multiple genes. In addition to insertions and deletions, there are several types of chromosomal rearrangements. Gene amplifications, translocations between chromosomes, inversions of chromosome sections, and mitotic recombination and gene conversion can all lead to rearrangements of the chromosomes and therefore a rearrangement of genes or parts of genes. If these events are large enough, they can be visualized under a microscope by someone trained in evaluating gross chromosome aberrations.

Mutations can occur spontaneously, for instance, by errors in DNA replication and DNA repair. Every organism has a spontaneous mutation rate, defined as the number of mutations per locus that are incurred either per cell division or, at the organism level, per generation. This rate of spontaneous mutation can vary across species, with some bacteriophage having rates as low as 1 in 10^9 loci. Higher organisms such as corn, *Drosophila*, and humans have rates of approximately 1 in 10^5–10^7 loci. In general, higher animals and plants have genes that spontaneously mutate at a much higher rate than those in microorganisms. It is unclear why this should be the case, but perhaps the basic difference is between diploid and haploid organisms. The former has no normal allele to offset the deleterious effects of a new mutation, while the latter are able to survive as long as a single normal gene can provide sufficient normal protein for cell survival. That is, new spontaneous mutations can be recessive in function if the cell is diploid, but in haploid cells they are effectively dominant and also generally lethal. It should be noted that in rare instances some new mutations may provide the organism a selective advantage and this is the basis for evolutionary modifications.

Since the demonstration by H. J. Muller in the early twentieth century, that X-rays can induce new mutations, the science of studying the ability of chemical and physical agents to cause genetic damage has grown rapidly. The field of genetic toxicology took shape in the 1960s and 1970s when scientists developed methods to assess mutation in a quantitative way using a wide variety of organisms from bacteria, yeast, *Drosophila*, *Neurospora*, plants, mammalian cells and various rodent species as model systems.

Penetrating ionizing radiation, including X-rays, gamma rays from radium and other radioactive atoms, and neutrons from nuclear reactors, has been widely studied. The first chemicals to be identified as mutagens include mustard gas, and various nitrogen or sulfur–mustard compounds.

A chemical that causes chromosome breakage is a clastogen. Clastogens are generally defined as those chemicals that can cause chromosome damage that can be seen using a microscope. If there is damage to the cellular machinery involved in cell division and chromosome pairs do not all separate normally into the two daughter cells, this nondisjunction will result in aneuploidy, an abnormal number of chromosomes in the daughter cells. Polyploidy results when there are full (or almost full) additional sets of a chromosome complement.

Following the metabolism of a chemical and its transport through the body and cells to reach the genetic material in the cell nucleus, there are a number of biological events that can lead to the induction of a mutation. First, the chemical may have some interaction with the DNA or the enzymes

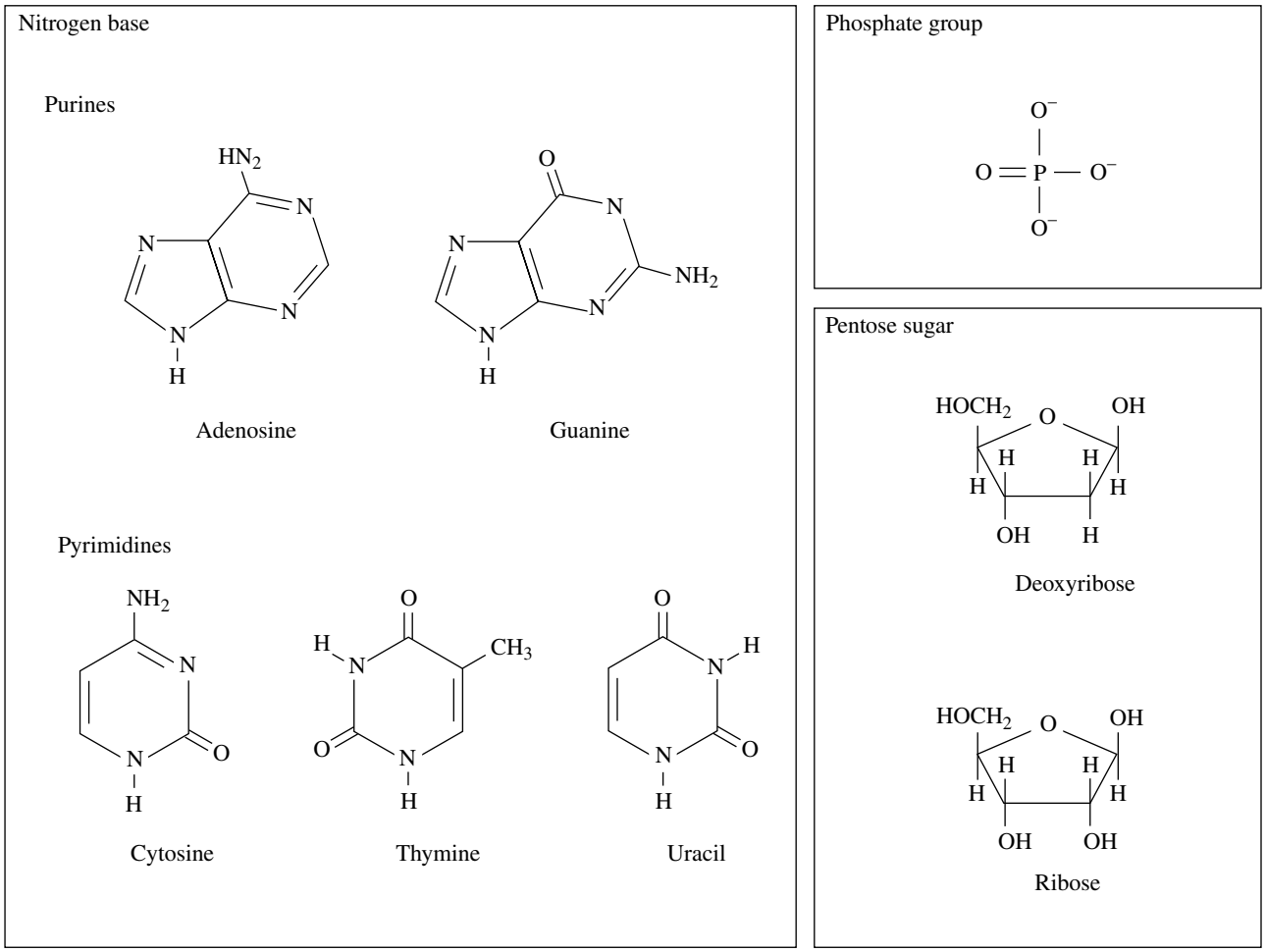


FIGURE 14.3 The purine and pyrimidine bases, phosphate group, and pentose sugars comprise DNA or RNA.

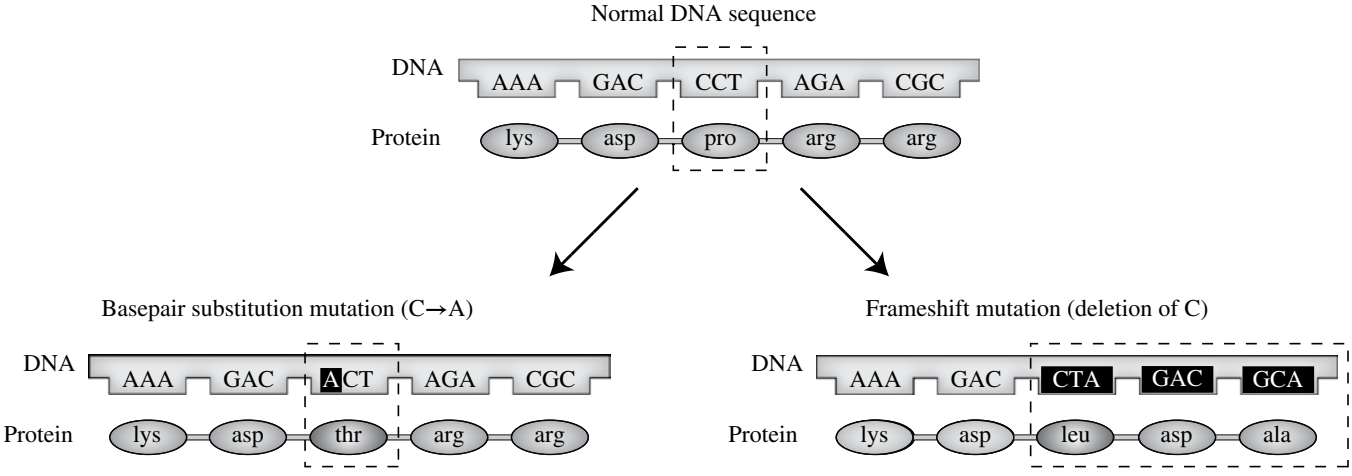


FIGURE 14.4 Single base changes can alter the normal DNA sequence and result in either a base pair substitution mutation or a frame shift mutation.

involved in DNA function. While this generally results in the formation of a DNA adduct (covalent binding of the chemical with the DNA), other interactions including intercalation within the DNA strand are also possible. The inhibition of topoisomerase is another mechanism by which some chemicals cause mutations.

Electrophilic chemicals that covalently bind to the DNA generally form DNA adducts containing at least a portion of the chemical. These chemical-specific DNA adducts can be identified and quantified. Such an analysis allows for a measure of the amount of primary DNA damage caused by the interaction of the chemical with the DNA. Some of these adducts can be relatively large as is the case with polycyclic aromatic hydrocarbons and heterocyclic amines. Simple alkylating agents generally form N7-methylguanine adducts. Both endogenous and exogenous reactive oxygen species form 8-oxyguanine adducts. Table 14.1 provides information about the types of DNA damage and examples of the chemicals/physical agents that cause damage and the specific mutation induced.

Chemical-induced DNA adducts can be repaired, misrepaired or the DNA may be processed with the adduct resulting in the incorporation of an incorrect base in the daughter DNA. DNA adducts are common. It is estimated that endogenous and exogenous chemicals can produce as many as a million molecular lesions per cell per day. Endogenous DNA damage is caused by various normal cellular processes including those that result from our reliance on oxygen for cellular metabolism and produce reactive oxidation species and misincorporation of DNA bases because of errors in replication. Exogenous DNA damage can be caused by a variety of chemical and physical agent exposures, some virtually unavoidable, including exposure to sunlight or artificial light. There are distinct differences between the types of mutations caused by endogenous and exogenous DNA damage and it is possible to provide insight into the cause of the mutational event by evaluating the DNA sequence. This pattern of mutational changes is called the DNA fingerprint or mutational spectrum.

In addition, alkylating agents can induce apurinic or apyrimidinic (AP) sites because the adducts destabilize the glycosidic bond. Hydrolysis of DNA bases can cause deamination, depyrimidination, or depurination. This can occur spontaneously or from chemical exposure. The resulting AP sites can result in opening of the ribose ring and cause single DNA strand breaks.

Some chemicals, particularly those containing two reactive sites, can form inter- and intrastrand cross-links in double-stranded DNA or DNA–protein cross-links. DNA cross-links can block replication and ultimately lead to cell death or to mutations in the daughter cells. For instance, carmustine (a cancer chemotherapeutic and alkylating agent) is particularly cytotoxic because it forms a cross-link between the N-1 of guanine and the N-3 of the opposing cytosine, which effectively blocks DNA synthesis and cellular replication. Cisplatin often causes intrastrand cross-links between adjacent N-7 guanines.

If a break occurs in a single strand of the DNA helix, it is called a single-strand break. These are among the most common DNA lesions and generally result from attack of the DNA by endogenous free radicals and alkylating agents. Single-strand breaks also are induced by ionizing radiation and can be a secondary consequence of the repair of DNA adducts caused by a number of mutagenic chemicals.

Double-strand breaks result from simultaneous breakage of the two complementary DNA strands. This can cause a

TABLE 14.1 Types of DNA Damage and Examples of Chemical/Physical Agents that Induce Specific Mutations

Type of DNA Damage	Causative Agent	Major Mutation
DNA hydrolysis products		
Deaminated bases	Spontaneous deamination of 5-methylcytosine	G:C→A:T transition
Apurinatic lesions	Alkylating agents	G:C→T:A transversion
Apyrimidinic lesions	Alkylating agents	A:T→ T:A transversion
DNA monoadducts		
8-oxoguanine	Reactive oxygen species	G:C→T:A transversion
Alkyl adducts	Alkylating agents (e.g., ENU), spontaneous methylation	Base pair substitution
Bulky adducts	PhiP, Benzo[a]pyrene, *N*-2-actylaminofluorene	Base pair substitution, frame shift
Cross-links		
Thymine dimmer	UV light, mitomycin C	Tandem mutation, base pair substitution
Interstrand crosslinks	Bifunctional alkylating agents (e.g., carmustine)	Frame shift, deletion, chromosome mutation
Intrastrand crosslinks		
DNA and protein crosslinks	Cisplatin, aldehydes	Basepair substitution Chromosome mutation
Single-strand breaks	Free radicals, ionizing radiation, UV, alkylating agents	Deletion, insertion, base pair substitution
Double-strand breaks	Ionizing radiation	LOH, deletion, translocation
Nondisjunction	Colcemid	Aneuploidy

physical separation of the chromosome and the loss of, in some cases, large amounts of genetic material. Such double-strand breaks will often result in cell death. Ionizing radiation and some alkylating agents induce double-strand breaks.

14.2 GENETIC DAMANGE AND ITS IMPACT ON HUMAN DISEASE

There are a large number of human diseases that are known to involve mutations inherited from an individual's parents. In some diseases, a single abnormal allele can cause the effect. These are referred to as dominant gene diseases. Examples of these include: Huntington's disease, chondrodystrophy, and hepatic porphyria. Recessive mutation diseases require that the individual inherit an abnormal gene from both parents; that is, two copies of the mutated gene are required for the disease to be expressed. Diseases like albinism, cystic fibrosis, xeroderma pigmentosum, sickle cell anemia, and type 1 diabetes mellitus are recessive diseases. In hemophilia, the presence of a normal allele will block the expression of the disease. Unfortunately, the gene for hemophilia is on the X chromosome and, therefore, males who inherit a single abnormal gene from their mothers will express the disease. Females, on the other hand, have to inherit abnormal alleles from both parents. This means that hemophilia is much more common in males than females. The same is true for color-blindness, the effects of which are, of course, much less severe. Other diseases are caused by abnormal genes having low penetrance, meaning that the severity of the disease can vary among individuals who carry the mutant allele. In addition, there are diseases with complex and multiple gene involvement so that the inheritance pattern is much less clear. Anencephaly, club foot, and spina bifida are examples of these diseases.

Abnormal chromosome number (aneuploidy), while generally lethal, also can result in human inherited diseases. The most well-known is probably Down syndrome in which the affected individual has three copies of chromosome 21 (trisomy). Edward's syndrome results from trisomy of chromosome 18. Other diseases resulting from an abnormal number of chromosomes include Klinefelter's syndrome (XXY) and Turner's syndrome (XO; only one sex chromosome). Cri-du-chat syndrome results from a partial deletion of chromosome 5. All of these chromosome abnormalities, while having severe health impacts, are compatible with the survival of the individual. There are various estimates (up to 60% during the first trimester) as to how many spontaneous abortions occur because the fetus has an abnormal chromosome complement that is not compatible with survival.

Somatic cell mutations, both inherited and newly acquired, are known to be involved in the etiology of cancer. The retinoblastoma gene (*RB*) is an extreme example of an inherited mutation involved in increased cancer risk. Individuals carrying an abnormal copy of the *RB* gene, that is, individuals who are heterozygous, need only a single mutation deleting the function of the normal allele to result in the formation of a retinoblastoma tumor. Another well-known example is the *BRCA* gene that dramatically increases the risk of breast cancer in women who carry a germ line mutation of this gene.

Generally, the etiology of cancer involves an increase in mutation and increased cell division. In some cases, the process is initiated by a newly induced mutation. In others, some biological event causes an increased rate of cell division and this increase in the number of cells carrying a specific cancer gene mutation is thus increased in frequency. Cancer genes include both oncogenes (a mutated form of a normal gene termed a proto-oncogene), which cause abnormal cell function, and tumor suppressor genes, which when mutated cause a loss of cellular control. Some of the best characterized oncogenes include *RAS*, *TRK*, *ERK*, *WNT*, and *MYC*. Probably the best known tumor suppressor gene is *P53*.

There is increasing evidence that somatic cell mutations may be involved in diseases other than cancer, including atherosclerosis, and cataracts.

14.3 GENETIC TOXICOLOGY TESTS

By the mid-1970s, there were a number of investigators who had developed tests for the detection of newly induced genetic damage. *A Handbook of Mutagenicity Test Procedures* was edited by B. J. Kilbey, M. Legator, W. Nichols, and C. Ramel and published in 1977 by the Elsevier Scientific Publishing Company. Chapters were written by the leaders in the newly evolving field of genetic toxicology. A wide variety of organisms, including bacteria, yeast, mouse cells, human fibroblasts, human peripheral blood lymphocytes, mice, wasps, *Neurospora*, onion root tips, and *Drosophila* were used for conducting genetic toxicology tests and most of these assays were included in the book. There were also a wide variety of endpoints, including DNA excision repair, gene mutation, mitotic recombination, micronuclei, gross chromosome aberrations, and sister chromatid exchange, described in the book. Additional chapters covered issues such as laboratory safety when working with mutagens, statistical considerations, and methods for obtaining epidemiological data for mutation risk assessment.

By the 1980s, regulatory agencies were beginning to discuss the various genetic toxicology tests that might be useful for regulatory decision making. While agencies took different strategies based on regional preferences or regulatory uses and mandates, some patterns emerged. Generally, agencies were interested in selecting a combination of assays that would cover the range of adverse genetic events, including point mutations and chromosomal events. In the United States, the Environmental Protection Agency (EPA)

selected a battery that included the Ames test, an *in vitro* mammalian assay (either the mouse lymphoma assay (MLA) using the thymidine kinase (*Tk*) gene or both a chromosome aberration assay and an assay using the hypoxanthine guanine ribosyltransferase (*Hprt*) gene) and an *in vivo* assay for chromosomal damage.

Many agencies/organizations now have Web sites which can be readily updated and easily accessed for the latest information. The manuscript by Michael Cimino (2006), listed in the suggested reading section, provides a useful summary.

The Organization for Economic Cooperation and Development (OECD) has published guidelines for the commonly used genetic toxicology assays. An expert workgroup was convened within the last couple of years and this group is currently revising the test guidelines.

Assays in the Standard Battery

As indicated earlier, the standard battery usually includes a bacterial assay, an *in vitro* mammalian cell assay capable of detecting chromosomal damage and an *in vivo* cytogenetic assay. All of the *in vitro* assays must be conducted both with and without exogenous metabolic activation because the cells used have limited capability to metabolize promutagens to their active metabolites. Generally, this activation is accomplished using a cofactor-supplemented S9 fraction prepared from the livers of rodents (generally rats) that have been treated with a chemical such as Aroclor 1254 or a combination of β-naphthoflavone and phenobarbital. The S9 used for the *in vitro* genetic toxicology assays is a mixture of microsomes and cytosol. It contains a variety of both Phase I and Phase II enzymes and includes cytochrome P450 enzymes, acetyltransferases, carboxyesterases, and other enzymes. Some investigators have developed methods whereby primary hepatocytes are cocultured with the mammalian cells used in genotoxicity assays, and it is possible to make S9 from a variety of species including humans. As a general rule, these latter activation methods are more common in research applications and are rarely used in routine *in vitro* genetic toxicology evaluations.

Bacterial Reverse Mutation Assays (OECD TG 471, FDA Redbook)

The most widely used genetic toxicology assay is the Ames test, which uses a set of specially designed strains of *Salmonella typhimurium* (Figure 14.5). The Ames test is commonly used as an initial screen to determine the mutagenic potential of chemicals. The assay detects point mutations using a set of tester strains containing preexisting mutations that require the presence of histidine in the medium in order to grow. Mutations that restore a functional sequence to the gene enable the bacteria to synthesize histidine. These mutations that change a bacterium back to function are called reversion mutations and the mutated bacteria, revertants.

Each of the tester strains has its own properties, having different types of mutations that must be reverted to histidine independence in order to grow into colonies that can be enumerated (Table 14.2). By using a set of these strains it is possible to detect and distinguish mutagens acting by various mechanisms. Some of the strains are reverted by base pair substitutions. Others contain frame shift mutations and therefore require frame shift mutations to restore function. Because bacteria have cell walls that are not as permeable as the cell membranes of mammalian cells, all

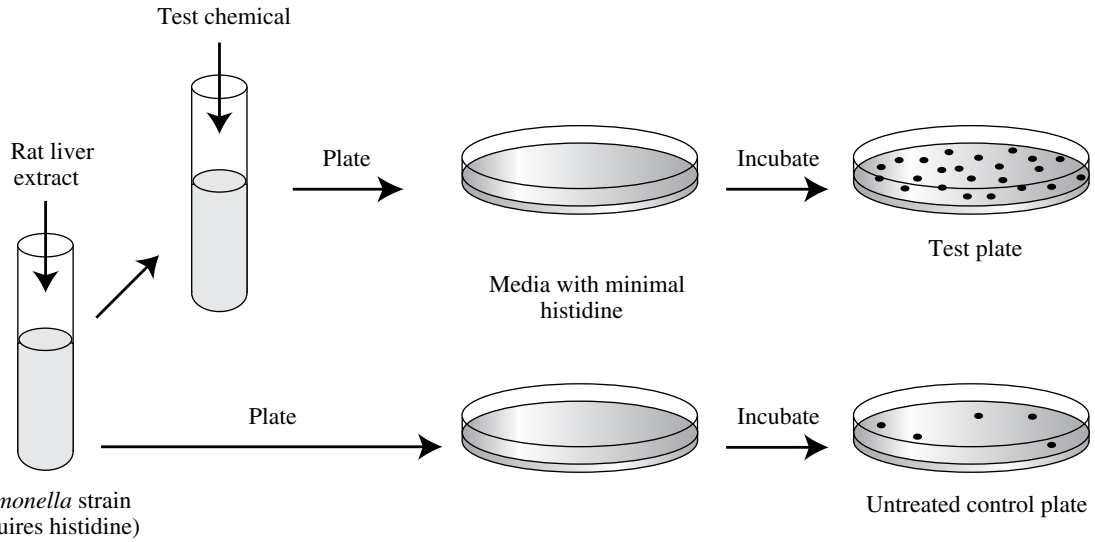


FIGURE 14.5 The Ames test is conducted by treating the various tester strains of *Salmonella* with the test chemical, plating in petri dishes and counting the number of revertants that can grow and form colonies.

TABLE 14.2 Commonly Used Ames Test Strains, the Types of Mutation Detected and Example Mutagens for Each Strain

Strain	Mutation Detected	Target Sequence	Mutagen
TA97	Frame shift	CCCCCC	4-NQO
TA98	Frame shift	CGCGCGCG	2-AF
TA100	Base pair substitution	GGG	BaP, MMS
TA102	Base pair substitution	TAA	H_2O_2, MMC
TA104	Base pair substitution	TAA	formaldehyde
TA1535	Base pair substitution	GGG	2-AA, NaN_3
TA1537	Frame shift	CCCCC	2-AA, AAC
TA1538	Frame shift	CGCGCGCG	2-NF

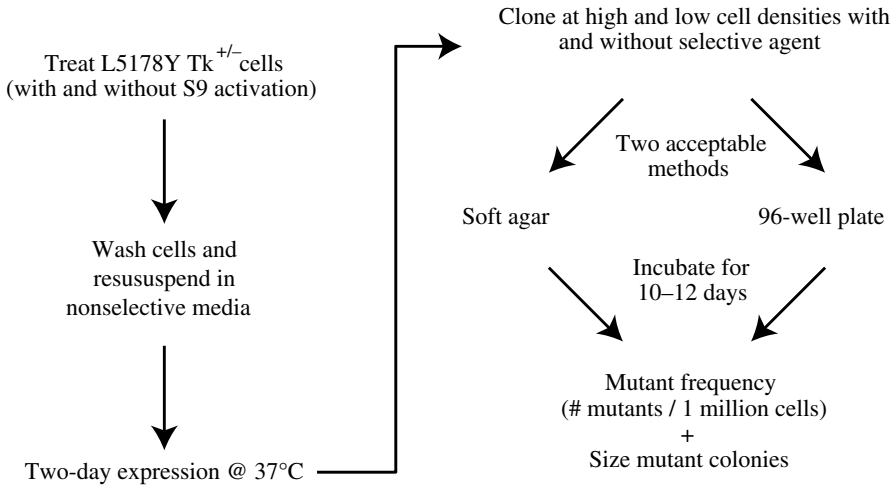


FIGURE 14.6 The MLA is conducted by treating cells with the test chemical, removing the chemical, resuspending the cells, and allowing the newly induced mutations to express. Following a 2-day expression period, the cells are cloned at a high cell density with the selective agent so that the mutants can grow and at a low cell density to determine plating efficiency. Following a 10- to 12-day incubation, the colonies are counted, mutant colony size determined, and mutant frequency calculated.

of the strains contain mutations that have damaged the cell wall in a way that allows test chemicals to enter the bacteria. In addition, some of the strains contain defects in DNA repair enzymes so that any initial DNA damage will not be repaired, but rather will result in revertants that can be grown into colonies which can be scored.

To conduct the assay, the bacteria are plated with a trace amount of histidine in petri dishes containing agar. There are three basic versions of the assay: the spot test, the plate incorporation test, and a version in which the chemical and bacteria are preincubated and then plated to enumerate the revertants. There are some strains that are engineered to express some metabolic capability, but in general the bacteria have little to no metabolic ability and it is necessary to use an exogenous metabolic system generally made from rat liver. Typically, the assay is conducted both with and without S9. This provides some information as to whether the chemical is itself capable of inducing mutation or if metabolism to active metabolites is a prerequisite for the chemical to induce mutations.

In addition to the Ames test, the *Escherichia coli* assay is sometimes used and is based on mutations in the tryptophane biosynthetic pathway. Both of these bacterial assays detect point mutations involving single or a very small number of base pairs and include substitutions (transitions and transversions), and the addition or deletion of one or a few base pairs. Therefore, a complete assessment of a chemical requires that additional tests be performed using mammalian cells and endpoints that detect both gene mutations and chromosomal events.

In Vitro Mammalian Cell Gene Mutation Assays (OECD TG 476, FDA CFSAN Redbook)

The MLA using the thymidine kinase gene (*Tk*) is the most widely used *in vitro* mutation assay (Figure 14.6). The next most widely used *in vitro* gene mutation assays utilize the *Hprt* locus. Because the *Hprt* gene is located on the X chromosome, it is possible to conduct the *Hprt* assay in most cell types. The *Tk* locus, on the other hand, is located on an autosome (chromosome 11 for mice, chromosome 17 for

humans) and so it is necessary to generate a *Tk* heterozygous line in order to detect mutation from $Tk^{+/-}$ to $Tk^{-/-}$.

In the early 1970s, Donald Clive and coworkers isolated a $Tk^{+/-}$ subline for the specific purpose of developing a mammalian cell mutation assay. This was accomplished using bromodeoxuridine to select clones that were deficient in the *Tk* enzyme. The growth of cells that have *Tk* enzyme is arrested, while cells that have no *Tk* enzyme are capable of growing in the presence of the selective drug and can form into colonies. Because the frequency of *Tk*-deficient cells is very low (on the order of 1×10^{-9}) in a population of $Tk^{+/+}$ cells, a very large number of cells were required to obtain a small number of $Tk^{-/-}$ clones. These *Tk*-deficient clones were expanded and allowed to grow into a population of cells. To obtain the $Tk^{+/-}$ clones required for the mutation assay, a selective medium containing methotrexate was used to yield clones that had a reversion of one Tk^{-} allele to *Tk* competency. Ultimately, a subline with the designation L5178Y $Tk^{+/-}$ 3.7.2C was developed and is now used routinely.

While the assay was originally developed using bromodeoxyuridine as the selective agent, a better selective agent was soon identified. Trifluorothymidine (TFT) was shown to provide for more rapid arrest of *Tk*-enzyme-containing cells and therefore to provide for better growth and more accurate enumeration of *Tk* mutant cells. Now, TFT is the only approved selective agent for use with the MLA.

The assay is conducted using a range of test chemical concentrations that cover a prescribed cytotoxicity range. The MLA was developed using a unique measure of cytotoxicity, the relative total growth (RTG), which takes into account the toxicity during the cell treatment and culture and also during the cloning phase of the assay. The recommended maximum cytotoxicity is between 20 and 10% RTG, which does not represent a "killing" of 80–90% of the cells. In fact, it is similar to the maximum cytotoxicity now recommended for the *in vitro* cytogenetic assays (see OCED TGs 473 and 487) for most chemicals.

Because newly mutated cells require time to become deficient in the *Tk* enzyme, a period of time called the phenotypic lag or expression time is required before the mutant frequency can be enumerated. For the *Tk* gene, it has been determined that 2 days are sufficient and optimal. It is important to note that longer expression times are not acceptable because the mutant frequency declines. Following the 2-day expression period, cells are plated at high cell density in medium containing TFT that will allow only *Tk* mutants to grow and also at a very low cell density to determine the plating efficiency of the cells. The mutant frequency is calculated based on the number of mutants that grow out of the population of cells plated in the TFT, corrected for the plating efficiency (i.e., the percentage of cells that can grow in the cloning plates). The plates are incubated for a sufficient time (about 10–12 days) to allow the colonies to attain sufficient size for counting. Mutant frequency is usually expressed as the number of mutants per million cells. The normal background (nontreated culture) mutant frequency is between 35 and 170 *Tk* mutants per million cells.

There are two different versions of the assay based on the method used to enumerate the mutants. The assay was originally developed using petri dishes and medium supplemented with agar to provide a semisoft medium in which round distinct colonies could form. Another version of the assay uses 96 well microwell plates and liquid (without agar) medium. Both versions of the assay are acceptable. It should be noted that the only difference in the two methods is how the mutants are selected and enumerated. The cell treatment and expression period are the same. A reference for the complete details for conducting the MLA can be found in the reading list.

The MLA has an interesting feature in that the *Tk* mutants form clones of two basic size classes. Cells in the large-colony mutants grow at the same rate as the parent cells, while the small-colony mutants grow much slower. A large body of research indicates that the induction of small-colony mutants is associated with exposure to chemicals that cause chromosome breakage (clastogens) and the induction of large-colony mutants is associated with exposure to chemicals that cause point mutations. This ability to detect both chromosomal mutations and gene mutations is a unique feature of the MLA. Using a combination of cytogenetic, molecular cytogenetic, and molecular genetic techniques to evaluate *Tk* mutant cells, it has been demonstrated that the MLA assay using the *Tk* gene is capable of detecting a wide spectrum of genetic events including point mutations, deletions, translocations, mitotic recombination/gene conversion, and nondisjunction (aneuploidy). Because of this ability to detect the array of genetic events associated with cancer and human genetic disease, the MLA is often listed as the preferred *in vitro* gene mutation assay in test batteries.

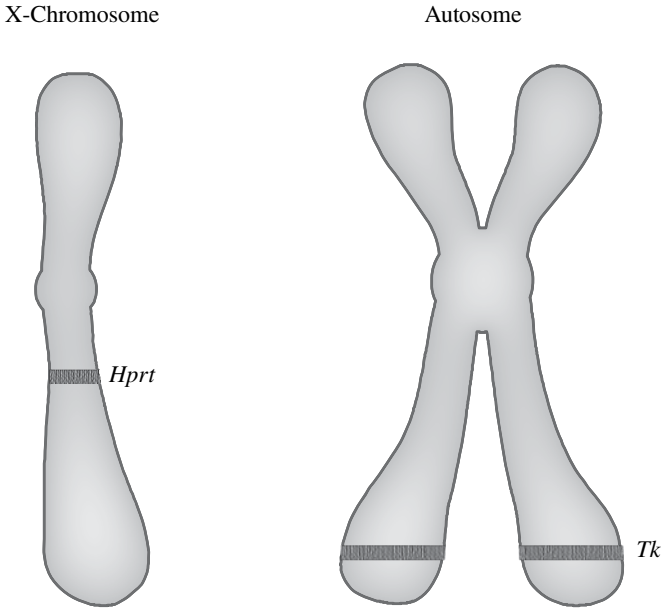
The TK6 assay was developed to be similar to the MLA in that it uses the human *TK* gene for mutation detection. The *TK* heterozygote used in this assay was obtained in a similar fashion as the one used in the MLA. The cell line used is a human lymphoblastoid line and is conducted using the same basic methods as the MLA. As with the MLA, the mutants can be classified as slow and normal growers and the assay detects the same range of genetic events as the MLA.

As indicated previously, the *Hprt* gene is located on the X chromosome and therefore mutation assays can be conducted using many different cell lines and primary cells, including a version of the assay that uses lymphocytes isolated from humans. The primary human lymphocyte assay is described in more detail in the section describing human biomonitoring. The two most widely used cell lines for routine *Hprt* mutation analysis are Chinese hamster ovary (CHO) cells and a line isolated from Chinese hamster lung cells (V79). The basic principles of the assay are similar to the MLA. The selective agent for mutant *Hprt* cells is 6-thioguanine (6-TG).

Cells that have *Hprt* enzyme activity are poisoned by 6-TG via the purine salvage pathway, while mutant cells can survive and form colonies. The expression period for *Hprt* is substantially longer than for *Tk* and can vary somewhat depending upon the cell line used. Generally, about 6–7 days is sufficient. Unlike the *Tk* mutants, most *Hprt* mutants grow at the same rate as *Hprt* competent cells and therefore their frequency does not decline with slightly longer expression times. As with the MLA, cells are treated with a range of chemical concentrations and newly induced mutants are allowed to express during a period of cell culture in nonselective medium. Cells then are plated both with and without 6-TG. After a suitable incubation time, the colonies are counted and the mutant frequency is derived from the number of 6-TG-resistant mutant clones and the number of viable cells plated.

The *Hprt* assays detect primarily gene mutations, that is, mutations that impact the *Hprt* gene. While some mutants are the result of large deletions, including some that greatly exceed the boundaries of the *Hprt* gene, the X-linked location prevents the detection of mitotic recombination/gene conversion. Chemicals that induce primarily chromosomal mutations induce only a small (or no) increase in the frequency of *Hprt* mutants. Figure 14.7 provides a summary of the genetic events detected by the X-linked *Hprt* gene and the autosomal *Tk* gene. In most test batteries, if an *Hprt* assay is used as the *in vitro* mutation assay, it is also necessary to conduct an additional assay to evaluate for chromosomal damage—either an *in vitro* chromosome aberration assay or an *in vitro* micronucleus assay.

Guidelines for performing the *in vitro* mammalian gene mutation assays can be found in OECD TG 476 and the FDA CFSAN Redbook. Because the MLA Expert Workgroup of the International Workshop for Genotoxicity Testing (IWGT) has reached consensus on a number of parameters for assay acceptance and for evaluating the data as positive or negative, a separate OECD guideline for the MLA is currently under development to include these new recommendations.



Base-pair changes	Base-pair changes
Frameshifts	Frameshifts
Deletions/insertions (up to megabase size)	Deletions/insertions
	Gross deletions (including multiple genes)
	Mitotic nondisjunction
	Mitotic recombination (LOH)
	Gene conversion
	Translocation

FIGURE 14.7 The types of mutations that can be induced and enumerated in mammalian cell gene mutation assays varies depending upon whether the gene resides on the X-chromosome or on an autosome.

This resulted in the need to revise OECD TG 476, to include only the *Hprt* assay. When finalized, both of the OECD TGs dealing with *in vitro* gene mutation assays will include new recommendations for determining the acceptability of individual experiments and the interpretation of data.

In Vitro Chromosome Aberration Assay (OECD TG 473 and FDA CFSAN Redbook)

Chromosome aberration assays use microscopic evaluation to detect visible alterations to the chromosomes. This includes simple chromosome breakage or breakage and reunion of one chromosome with another chromosome. Events affecting a single chromatid of the chromosome can also be scored. These events are defined as chromosome and chromatid aberrations. It is also possible to identify polyploidy (alterations in the number of sets of chromosomes) and small changes in chromosome number (aneuploidy). Different chemicals interact to different degrees with the chromatids or chromosomes and thus induce different types of damage. Some chemicals cause both chromosome damage and alterations in chromosome numbers. A small number of chemicals induce only alterations in chromosome number. Chemicals that alter chromosome number are said to cause aneuploidy and are called aneugens.

The assay is conducted using either cultures of established cell lines (CHO, V79 or Chinese hamster lung (CHL)) or primary cell cultures (human lymphocytes). The CHO cell line is currently the most frequently used. Cell cultures are exposed to the test substance with and without S9 exogeneous activation. After a period of culture optimized to maximize the number of first division metaphases, the cells are treated with colcemid to inhibit cells from completing mitosis, harvested using a mixture of acetic acid and methanol to "fix" the cells, and then fixed cells are dropped onto slides and stained with a dye that allows easy visualization of the metaphase spreads. Microscopic analysis is used to score the presence of chromosome aberrations. Generally, the analysis includes an enumeration of breaks, fragments, and translocations as well as alterations in chromosome number and any cells that are polyploid. It should be noted that it is difficult to adequately detect aneuploidy in the standard screening test. Data from treated cultures are compared with the negative and/or solvent control to determine if the chemical caused a significant increase in chromosomal damage. It is important that positive control chemicals be used in each assay to assure that the assay is working properly and that the individual scoring the slides is conducting a proper analysis. Generally, for this analysis, the slides are coded in a way that prevents the scorer from knowing the treatment conditions. This blind scoring is specific for cytogenetic assays because there is a significant degree of subjectivity in the analysis.

The concentrations used in current protocols for cytogenetic tests are selected to cover a specific range of cytotoxicity to the cells in culture. It should be noted that much of the earlier data for the *in vitro* cytogenetic assays was obtained with little concern for the degree of cytotoxicity. Often the top concentration evaluated showed a significant reduction in the number of cells that were viable enough to reach metaphase. This should be taken into account when weight-of-the-evidence evaluations are performed using older *in vitro* cytogenetic studies. Because there is often no reported measure of cytotoxicity, it is not possible to assess whether a reported positive result occurred using concentrations now considered acceptable, or concentrations that resulted in nonphysiological conditions and now considered biologically irrelevant.

OECD TG 473 for conducting the *in vitro* chromosome aberration test was revised and provisionally accepted in April 2014. This revised TG includes new recommendations for individual experiment acceptability and data interpretation as well as recommendations for assessing and limiting cytotoxicity. The measure to be used for cytotoxicity is consistent with that recommended for the *in vitro* micronucleus assay (TG 487; first adopted in 2010).

In Vitro Micronucleus Assay (OECD TG 487)

The *in vitro* micronucleus assay was adapted from the *in vivo* version of the assay by Fenech and Morley in 1986. They were the first to incorporate cytochalasin B to block cytokinesis. Micronuclei are formed when acetic fragments of chromosomes or whole chromosomes do not segregate normally during cell division into the daughter cells and these chromosome structures are enclosed by a nuclear membrane to form a satellite structure that appears as a very small nucleus (a micronucleus) (Figure 14.8).

As with the other *in vitro* genetic toxicology assays, cell lines or mitogen stimulated primary lymphocytes are treated both with and without S9. A full dose range of chemical concentrations should be used. A new OECD guideline for the micronucleus assay was recently approved (TG487, 2010). This guideline introduces and recommends two "new" measures of cytotoxicity to be used when the assay is conducted with cell lines. Both of these measures use the cell counts obtained at the time of treatment and approximately 24 h after treatment. The relative increase in cell count (RICC) is calculated based on comparing the relative increase in cell count for the treated culture versus the control culture. The relative population doubling (RPD) is calculated based on a comparison between the number of population doublings of the treated culture versus those of the control culture. It is important to remember that these two measures may give a different indication as to the level of cytoxicity induced by the treatment. They are also both dependent upon the length of the normal cell population doubling of the particular cell line used. When primary lymphocytes are used, it is not possible to use these measures because a variable

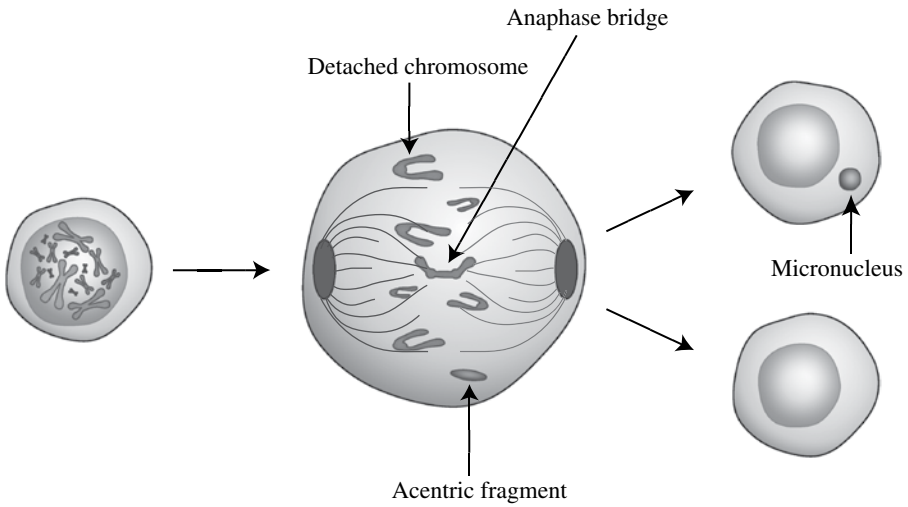


FIGURE 14.8 Micronuclei are formed in mammalian cells when pieces of chromosomes or whole chromosomes do not attach to the mitotic spindle and then fail to be incorporated into the nucleus of a daughter cell. These chromosome fragments or nonsegregated whole chromosomes appear in a daughter cell as a small nucleus (micronucleus).

proportion of the target cell population is mitogenically stimulated to divide. In this case, a cytokinesis blocking agent (i.e., cytochalasin B) is generally used and cytotoxicity is estimated from the relative frequency of binucleate cells in the treated cultures.

Also, in the OECD TG 487 guideline, the recommendation is made that the limit level of cytotoxicity be 55 ± 5% of the vehicle control (or 45 ± 5% RICC or RPD). As indicated previously, the OECD genetic toxicology guidelines are currently being revised. TG 487 also has undergone some revision and a revised version was provisionally approved in April 2014. The TG 487 revision includes some new recommendations for individual experiment acceptability and data interpretation.

Micronuclei are generally scored in the first interphase following exposure. This is the best time to observe the lagging chromosomes and chromosome fragments as replication of damaged DNA is necessary to form micronuclei. Because cytochalasin B prevents the cells from actually dividing into two cells, it "holds" the cells after the formation of the two nuclei and therefore maximizes the number of cells completing a single cell division and available for scoring. That is, cells that have completed cell division but not separated into two daughter cells will have two nuclei which can be readily identified. Only those cells are evaluated for micronuclei. While it is not necessary to use cytochalasin B when using cell lines, as all of the cells are assumed to be equally capable of cell division, it is recommended for use with primary lymphocytes where variable proportions of the cells may divide. Although methods utilizing cytochalasin B appear to be the surest way of quantifying micronuclei, cytochalasin B may occasionally affect micronucleus formation.

In Vivo Chromosome Aberration Assay (OECD TG 475)

The third component of the recommended "standard" battery is an *in vivo* test for chromosomal damage. Historically, this was often the *in vivo* chromosome aberration assay, but more recently the *in vivo* micronucleus assay (see OECD TG 474) has gained in popularity primarily because of its ease of use. Generally, the bone marrow has been used for scoring aberrations using a microscope and imploying the same criteria as used for the *in vitro* aberration assay. Slides are prepared and stained and metaphase cells are evaluated. It is necessary to sacrifice the animal in order to obtain the cells from the bone marrow.

OECD TG 475 has been recently revised and the revision was provisionally accepted in April 2014.

In Vivo Micronucleus Assay (OECD TG 474)

The *in vivo* rodent micronucleus assay is used to detect cytogenetic damage in either the bone marrow or peripheral blood cells. Endpoints detected in the *in vivo* micronucleus assay are chromosome fragments or whole chromosomes that were not included in the cell nucleus at cell division. Micronuclei containing these fragments or whole chromosomes can be visualized microscopically or by methods using a flow cytometer. In mammals, the main nucleus is extruded during the early development of erythrocytes. This makes it particularly easy to identify micronuclei in these blood cells. An increase in the number of cells containing micronuclei is used as a means to identify chemicals that can cause chromosomal (but not point mutational) damage. Because whole chromosomes have a centromere (or kinetochore) that can be stained, it is possible to distinguish between cells having chromosome fragments and those with whole chromosomes

by staining for kinetochores. Micronucleus analysis (both *in vitro* and *in vivo*) is the easiest method to identify those few chemicals that specifically induce aneuploidy but little or no other types of genetic damage.

Rodent bone marrow has historically been used for the micornucleus assay. More recently, particularly with the development of flow cytometric methods to identify cells with micronuclei, peripheral red blood cells have increasingly been used. Rodents are typically exposed to the test agent either by intraperitoneal injection or by an exposure route relevant to the particular chemical being evaluated. There are two different treatment schedules recommended in OECD TG 474. If animals are treated once with the test substance, then samples are taken from the bone marrow at least twice between 24 and 48 h posttreatment. Samples from the peripheral blood are also taken at least twice starting no sooner than 36 h and at appropriate additional intervals not exceeding 72 h. When two or more daily treatments are used, bone-marrow samples should be collected once between 18 and 24 h after the last treatment and peripheral blood samples should be collected once between 36 and 48 h after the last treatment.

One advantage of the peripheral red blood cell analysis is that blood can be taken from the tail vein and several sampling times can be used without the need to sacrifice the animal. The bone marrow is taken from the femurs or tibias. In both cases, for microscopic analysis, the cells are spread onto slides and stained to better visualize the cells and micronuclei. Automated methods and flow cytometry greatly increase the throughput of the analysis, lessen variability because a large number of cells can be scored, and thus improve the statistical power of the analysis. It is also much more feasible to perform the peripheral blood assay with flow cytometry on rats and other species whose spleens remove micronucleated erythrocytes. This technology has the throughput to quickly evaluate the small fraction of newly synthesized blood cells prior to the loss of micronuclei in the spleen.

OECD TG 474 has been recently revised and the revision provisionally accepted in April 2014.

In Vivo Somatic Cell Mutation Assays (OECD TG 488)

Although not a part of the routinely used screening battery, there are several *in vivo* mutation assays that can be conducted and provide useful information for hazard identification, for mode-of-action, and other risk characterizations. These include transgenic mutation assays and assays using peripheral blood lymphocytes and the *Tk*, *Aprt*, or *Hprt* gene.

Endogenous gene assays are limited to those genes for which there is a single copy (or engineered to contain a single copy) and a means to apply selective pressure that will allow only mutant cells to grow. These assays generally are also limited to lymphocytes. With the goal of being able to assess mutations in a variety of target tissues, several transgenic mutation assays have been developed and a recently approved OECD guideline provides guidance in using these tests. The most widely used include the Big Blue® rodent (mouse and rat; Figure 14.9) and MutaMouse™. Both of these assays employ transgenes inserted into the rodent

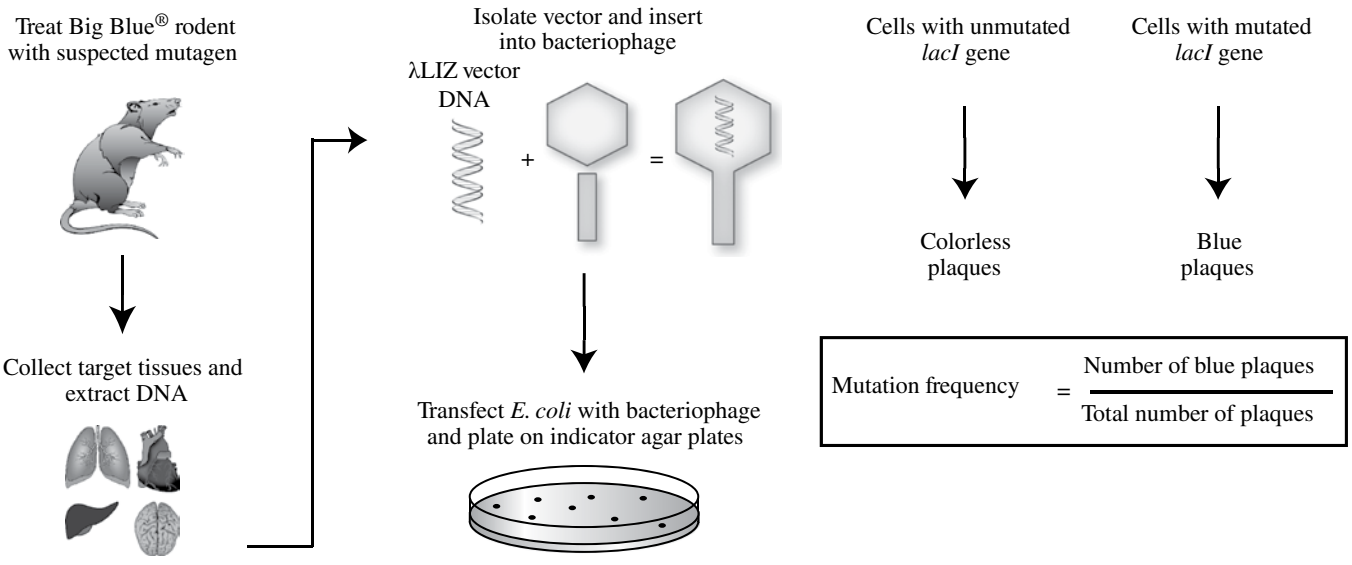


FIGURE 14.9 The Big Blue rodent assay is conducted by treating the transgenic animal by an appropriate route and duration (generally 28 days). Following sacrifice, the tissues are removed and flash frozen. DNA is isolated from the tissue of interest and the vector containing the transgene separated from the rest of the DNA so that it can be packaged into a bacteriophage, which is then transfected into *E. coli*. For the *lacI* endpoint, unmutated DNA forms a colorless plaque and mutated DNA forms a blue plaque. Mutant frequency is calculated based on the number of blue plaques among the total number of plaques.

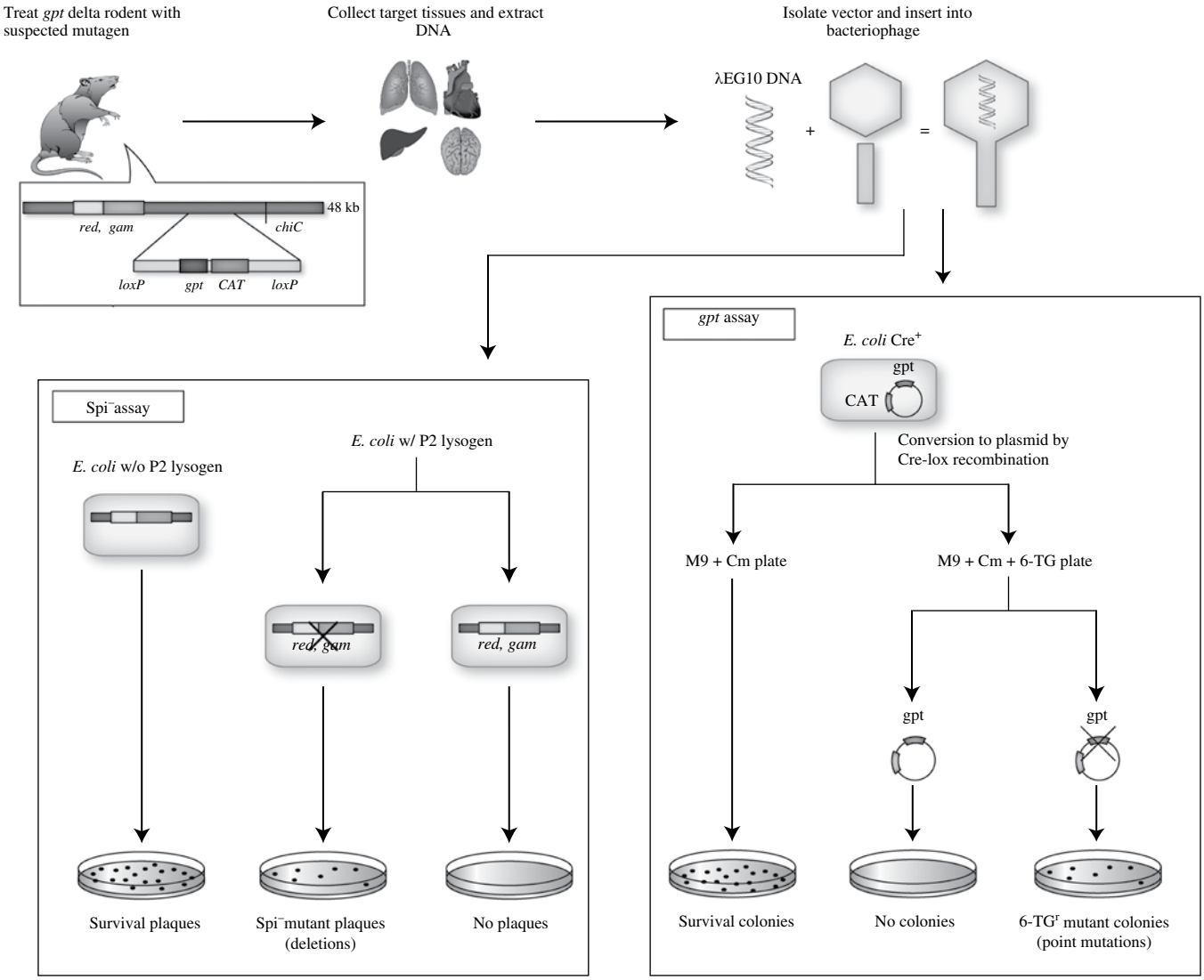


FIGURE 14.10 The *gpt* delta assay is conducted by treating the transgenic animal using an appropriate route and duration. Following sacrifice, the tissues are removed and flash frozen. DNA is isolated from the tissue of interest and the vector containing the transgene separated from the rest of the DNA so that it can be packaged into a bacteriophage, which is then transfected into *E. coli*. Bacteria containing mutations in the *gpt* locus are selected using 6-thioguanine, while Spi selection is used to detect deletions.

genome, which allow for the detection of point mutations and small deletions. Japanese researchers have developed another transgenic system, the *gpt* delta assay, which provides a means to detect both point mutations and up to approximately 10 kb deletions. While Big Blue and MutaMouse have been used by researchers in a number of countries, the *gpt* delta assay has primarily been used in Japan but due to its ability to detect a broader range of genetic events, it is beginning to be used in other countries (Figure 14.10). Because all of these transgenic models can be used in any tissue/organ from which DNA can be isolated, these models are particularly conducive to studies designed to assess mode-of-action (see Section "Mode-of-action for Cancer").

For all three transgenic models, transgenic animals have been engineered so that all of the cells of the animal contain a foreign target gene. The reporter transgene is of bacterial or viral origin and is located in a shuttle vector that is a derivative of a bacteriophage. The vector can be recovered from DNA isolated from virtually any tissue and then grown in an appropriate bacterial host. The transgenic animals are exposed to the test substance by an appropriate route. While the literature contains a wide variety of treatment protocols and primarily studies using acute exposures, the current recommendation is for a 28-day treatment regime followed by a 3-day expression period. The rodents are then sacrificed and the tissue(s) of interest harvested. Tissue can be flash frozen and maintained at −80°C for long periods of time. DNA is

isolated from the tissues and single copies of the vector DNA are excised from the high-molecular-weight DNA. The single vectors are packaged into infectious virus particles. *Escherichia coli* host cells are infected, plated, and incubated. Plaques or colonies (depending upon the endpoint) containing individual transgene vectors become visible on the plates and are counted. Mutant frequencies are determined by comparing the number of mutant plaques/colonies to the total number of plaques/colonies screened.

The Big Blue and MutaMouse animals are available commercially. The Big Blue models use the *lacI* gene (which has 1080 base pairs) or the *cII* gene (which has 294 base pairs). The *lacI* gene assay identifies mutants by their color and requires the plating of a large number of host cells on a large number of bacterial plates. The *cII* gene method is a positive selection method in which only mutant colonies grow on the selection plates (much as is done for the *in vitro* gene mutation assays). Because it is much easier to perform, most recent studies use *cII*. MutaMouse uses *lacZ* (which has 3100 base pairs) or *cII* as reporter genes and assays mutations in both genes using positive selection protocols.

In Vivo Germ Cell Genotoxicity Assays

While genetic toxicology assays are generally conducted to assess possible effects in somatic cells, there are also several tests that were developed specifically for germ cell damage and/or heritable mutation. The U.S. EPA Guideline for Heritable Mutation provides information on the use of these assays and there are OECD test guidelines that provide information on the conduct of the assays. As with the other OECD genetic toxicology test guidelines, these are currently undergoing revision. Although there are a number of germ cell assays available that have been used in the past, currently the transgenic rodent assays described in TG 488 are considered to be the most practical way of assessing germ cell mutation.

14.4 CHROMOSOME TRANSLOCATIONS IN GERM CELLS

Chromosome translocations (OECD TG 483) can be assessed in spermatocytes of treated male mice and in the spermatocytes of male offspring. While the test using treated males evaluates whether there is damage to the germ cells of the exposed individuals, the test using the male offspring of the treated males is a direct evaluation of the heritability of newly induced germ cell gene damage. For this assay, male mice are treated and the spermatocytes are evaluated 50–100 days following treatment. This delay in analysis allows for the damaged germ cells to progress to the spermatocyte stage. The collected spermatocytes are placed on slides and cells in metaphase are evaluated for chromosomal translocations.

The evaluation of the offspring, called the F_1 heritable translocation test (OECD TG 485), assesses the ability of the damaged sperm from the male parent to actually fertilize an egg and develop into a viable adult. To perform the test, the treated males are mated and the dividing testicular cells of the F_1 offspring are evaluated for the presence of translocations. It is possible to identify the germ cell maturation stage (spermatozoa, mature spermatid, young spermatid, spermatocyte, and spermatogonia) in which the damage occurred by mating the treated males at various times posttreatment. Chemicals that can cause translocations in the F_1 offspring are of particular concern because this test clearly shows that the newly induced damage is heritable and can be passed to subsequent generations.

14.5 THE MOUSE-SPECIFIC LOCUS TEST

Developed by William and Liane Russell at the Oak Ridge National Laboratory, the mouse-specific locus test detects chemically induced mutations that are inherited by the offspring of treated mice. The test requires specially bred mice that are homozygous mutants for a particular set of recessive genes that have specific phenotypes. Treated or untreated wild-type mice are bred to mice carrying these recessive genes. The offspring of this mating are then evaluated for various mutant phenotypes. The majority of the offspring will be heterozygous for the genes in question and will not show the phenotype. Because the expression of the phenotype requires two mutant copies, only those offspring with a new mutation in the specific gene will demonstrate the phenotype. Because of the number of animals and amount of resources needed for this test, the specific locus test is seldom performed today, and the OECD guideline covering the test recently has been retired.

14.6 THE DOMINANT LETHAL ASSAY IN MALE MICE (OECD TG 478)

Dominant lethal mutations will be expressed in any zygote that carries the mutation and zygotes with dominant lethal mutations will not form a viable offspring, thus reducing litter size. Also, because these mutations are lethal, any zygote carrying the mutation must result from a newly mutated gamete (either egg or sperm). This test does not require a specific strain of mice. But it is desirable to use mice that normally have a large litter size, so it is best that F_1 hybrids or outbred strains be used.

Generally males are treated with the test agent, but the test can also be conduced using treated females. The treated mice are mated by caging the male mice for 1 week with three virgin females that are 8–10 weeks old. The presence of a vaginal plug is used to confirm the date of the mating

and the initiation of the pregnancy. Seventeen days after the mating, the females are sacrificed and the uterus dissected and evaluated for normally developing fetuses and fetuses that have died. By comparing the number of dead fetuses in the treated and untreated females, it can be determined if the treatment caused an increase in fetal lethality, that is, dominant lethal mutations.

Human Biomonitoring

It is possible to use genetic endpoints to monitor humans exposed to chemicals via, for instance, occupational exposure. The science of human biomonitoring has evolved significantly over the last few decades. The basic paradigm for human monitoring is shown in Figure 14.11. As noted, there is a continuum of events that can be evaluated and measured

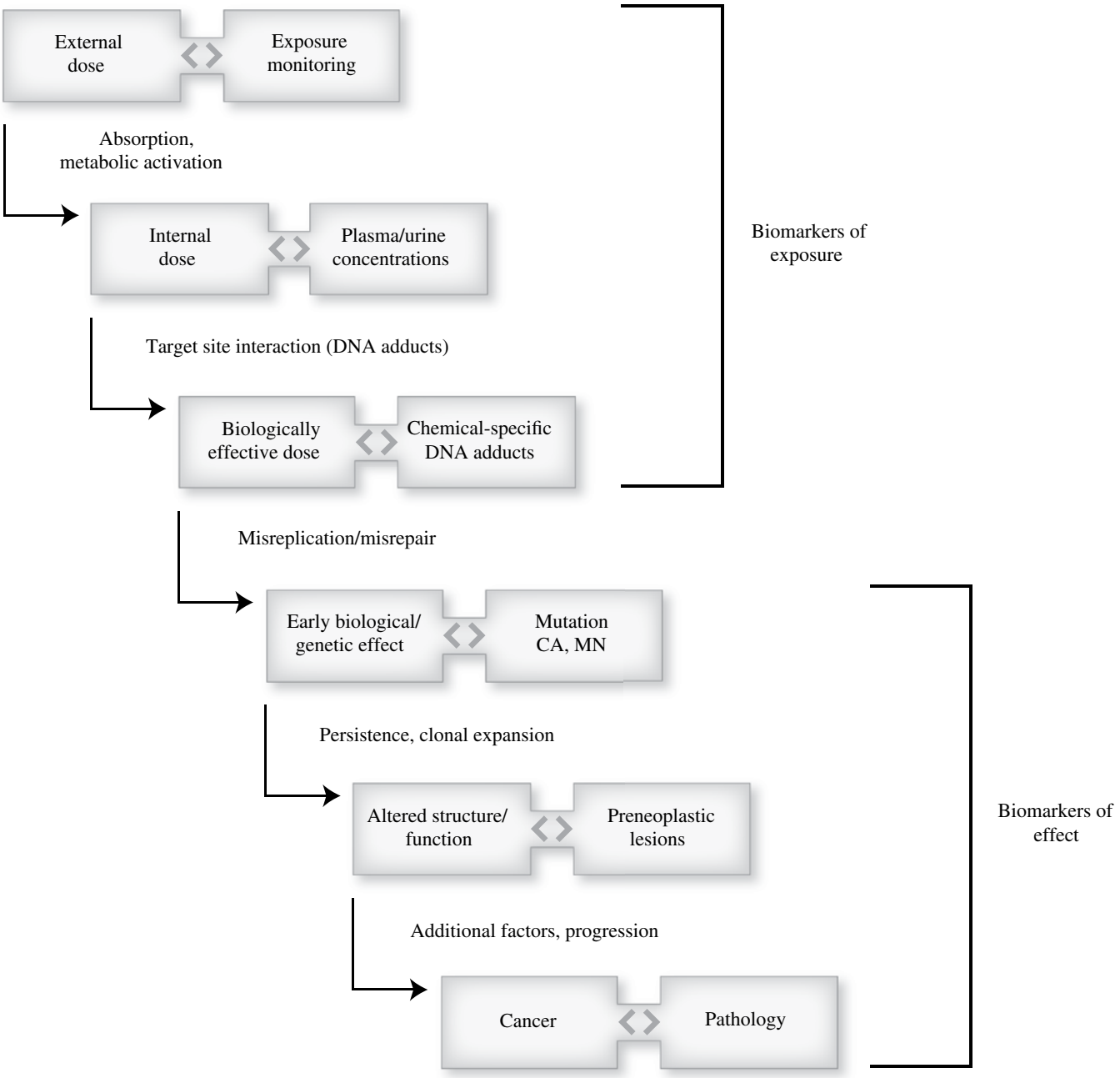


FIGURE 14.11 Basic paradigm for human monitoring. A comprehensive analysis of human exposure to environmental agents includes an analysis of exposure, human dose, target tissue dose, early biological effects, and, in the case of cancer, an assessment of the final cancer pathology. There are various biomarkers that can be used to assess exposure. Genetic toxicology biomarkers (gene mutation, micronucleus (MN) and chromosome aberration analysis (CA)) serve as surrogate markers for the early biological effects that occur following the transport of the chemical to a target tissue.

beginning with the chemical exposure, moving through the monitoring of chemicals and their transport in the body, biochemical or biological effects from the exposure, and ultimately the evaluation of human disease. Genetic endpoints generally are used as biomarkers for cancer.

The exposure evaluation for human biomonitoring is often the most complex assessment because it is necessary to consider all routes of exposure and all potential sources of exposure as well as the presence of multiple chemicals. It is this latter fact that makes human biomonitoring with the goal of associating particular exposures to adverse health outcomes extremely challenging.

Blood and urine are by far the most used matrices to assess the presence of chemicals in the body. Chemical analysis can detect a variety of chemicals such as metals, phthalates, pesticides, perfluorinated chemicals, polyaromatic hydrocarbons (PAHs), dioxins, environmental tobacco smoke, and aromatic amines in blood or urine. Protein adducts, in particular hemoglobin adducts, and DNA adducts, can be used as measures of exposure to chemicals and to determine the ability of these chemicals to interact with cellular material. DNA adducts provide an indication that the chemical can actually bind to DNA and therefore provides a strong indication that the chemical may be capable of causing mutational damage. DNA adduct analysis can be conducted using ^{32}P-postlabeling, which is a sensitive but not particularly specific method. Immunological methods provide sensitivity and some degree of specificity. Analytical chemical methods are available or can be developed that will provide specific information as to the type of chemical moiety bound to the DNA and are very useful to document specific exposures.

Genetic biomarkers that have been applied in human monitoring and other human studies include cytogenetic endpoints (chromosome aberration and sister chromatid exchange), gene mutation (*HPRT* and GPA (glycophorin A)), and also various DNA breakage assays including the Comet assay. Chromosomal endpoints have been evaluated in a European collaborative project (CancerRiskBiomarkers). European cohorts of almost 22,000 individuals have provided extensive information concerning the utility of chromosome aberration and sister chromatid exchange analysis in peripheral blood lymphocytes to predict human cancer risk. From this large study, it has been concluded that a high level of chromosome aberrations is associated with an increased risk of cancer, while no association was found for sister-chromatid exchange frequency and cancer. It appears that both chromatid- and chromosome-type aberrations predict cancer; however, there is some evidence that chromosome-type events have a more pronounced predictive value than chromatid-type events. A recently reported study of a large international cohort of 6718 people from 10 countries showed a significant association between micronucleus frequency in healthy individuals and cancer risk.

14.7 REGULATORY USE OF GENETIC TOXICOLOGY DATA

There are a number of different regulatory uses for genetic toxicology data. While some agencies, including the U.S. EPA, consider heritable mutation to be an important endpoint and have a guideline for conducting heritable mutation evaluation, the primary uses of the data are in cancer assessment decisions. The major regulatory uses include the identification of hazard and, for some carcinogens, mode-of-action evaluation for the induction of tumors.

The amount of information that is available for a regulatory decision can vary widely.

The preclinical safety evaluation of a new human drug includes conducting the test battery recommended by the International Committee for Harmonization (ICH). The issue under consideration is whether the drug is "safe" enough to be given to healthy human volunteers. Because genetic toxicology testing occurs relatively early in the safety assessment, the full set of studies that would be conducted for the ultimate approval decision has not yet been conducted. Therefore, genetic toxicology plays a key role in this stage of the process.

In some situations, such as for veterinary drugs and other chemicals with a possible residue in human food, positive responses from the genetic toxicology battery can trigger a request for a full cancer bioassay. In other situations such as for the registration of pesticides, the U.S. EPA receives a submission that includes a full set of required studies. In this case, the genetic toxicology data can be used as a part of the weight of evidence to determine whether the chemical is actually a carcinogen and/or likely to be carcinogenic to humans. The data also can be used as a part of the mode-of-action evaluation as to whether the tumors are formed through a mutagenic mode-of-action.

For many regulatory decisions, the amount of information provided to the agency or available in the literature can be very minimal. Often only basic genetic toxicology data will be available. For instance, for dietary supplements a regulatory decision must be made on the available information with no option for requesting further testing.

For large production volume chemicals and other chemicals that undergo a full risk assessment, including perhaps a quantitative assessment for cancer and noncancer adverse health outcomes, genetic toxicology data will be a part of the overall assessment and can play a role both in the determination as to whether the chemical is likely to cause cancer in people and if so, the mode-of-action for the induction of that cancer. The U.S. EPA IRIS program assimilates all the available information and conducts various quantitative assessments. WHO and IARC are international agencies that conduct evaluations of individual chemicals and in some cases classes of chemicals and also for environmental exposures to complex mixtures such as cigarette smoke.

Interpretation of *In Vitro* Genotoxicity Assays

There are many advantages to the standard *in vitro* genotoxicity assays. All of the *in vitro* assays are relatively easy to conduct and provide data in a few days to a few weeks. In addition, the assays are relatively inexpensive and do not involve the use of live animals, so they avoid all of the issues associated with the appropriate use of animals. There are also large databases for the standard assays. It is, however, important to note that the standards and acceptance criteria for assays and the interpretation of data, particularly for the *in vitro* mammalian assays, have evolved over time. Hence, one should not generally accept the positive/negative calls of the original authors without reviewing the data to assure that the call would stand based on current practice.

The biggest problems with the *in vitro* assays revolve around their inability to fully integrate the *in vivo* absorption, distribution, metabolism, and excretion/elimination, which can result in specific tissue (target tissue) susceptibility. Many promutagens require metabolic activation which occurs *in vivo* and in different tissues to varying degrees. Promutagens and metabolites also are transported through the body and can act differently in different tissues. This deficiency in metabolism is generally dealt with by adding a crude rat liver S9 mix and this provides many of the enzymes required. It is always, however, important to bear this difference in mind when conducting an overall weight-of-the-evidence assessment for a particular chemical.

Interpretation of *In Vivo* Genotoxicity Assays

It is widely accepted that the data obtained using *in vivo* assays are more relevant for predicting risks to humans because whole-animal exposures will do a better job of modeling the metabolism and pharmacokinetics that occur *in vivo* but cannot be duplicated *in vitro*. Unfortunately, *in vivo* systems are not available to fully assess the entire array of genetic damage that can be detected in the battery of *in vitro* tests. It is not unusual for a chemical to give a positive response in an *in vitro* assay yet be negative in the routinely conducted *in vivo* cytogenetic assays either in bone marrow or blood. It is not appropriate to totally dismiss positive *in vitro* findings in the face of negative *in vivo* data. A full evaluation should be conducted to understand the significance of the *in vitro* response in the absence of a positive *in vivo* response. Some mutagens induce mutations in specific tissues and are therefore missed in the routine *in vivo* tests which generally use blood cells. It is also possible that mutagens can be missed in the *in vitro* assays because of this tissue specificity. For instance, the drug tamoxifen, used to reduce the risk of breast cancer in high-risk women or to reduce the risk of breast cancer reoccurrence in breast cancer patients, induces endometrial cancer in women and both liver and endometrial tumors in rats. Yet, it is negative in the Ames test and some other *in vitro* tests. It is, however, positive when evaluated *in vivo* using the target tissue (liver) in a transgenic rat mutation assay.

Weight-of-the-Evidence Evaluation for Mutagenicity

There are a very large number of genotoxicity assays that have been used over the last several decades, some of which have very small databases, that is, the data for a particular chemical might come from a variety of assays, including from assays that have fallen by the wayside. This can make it difficult to decide whether or not a given chemical is a mutagen. Unfortunately, one will find that often the summary of the genetic toxicology data in even "comprehensive" data reviews and official documents from regulatory agencies are only very superficial. In some cases, the "weight-of-the-evidence" is merely a simple assessment as to whether there are more "positive" or more "negative" results. On the other hand, the assessor may conclude that a mixture of positive and negative results is "inconclusive." Neither approach is appropriate. To conduct a proper evaluation, it must be appreciated that different assays detect different types of genetic damage (Table 14.3). For instance, a negative Ames test and a positive *in vitro* assay for cytogenetics or the MLA can be a perfectly valid set of data and should be interpreted to indicate that the chemical does not induce point mutations, but is rather a clastogen, inducing chromosomal damage.

In conducting a weight-of-the-evidence evaluation it is critical that the reviewer actually examine the original data for the chemical under consideration. Unfortunately, over the years there have been a number of changes in how assays, particularly the "standard" assays, are conducted and the data evaluated. It is important that data be evaluated by

TABLE 14.3 Types of Genetic Damage Detected in the Different Standard Genetic Toxicology Assays

Mutation	Ames	Micronucleus	Chromosome Aberration	Mouse Lymphoma Assay
Insertion/deletion	Yes	No	No	Yes
Allele loss	No	No	No	Yes
Multilocus mutation	No	No	No	Yes
Mitotic recombination	No	No	No	Yes
Large chromosomal damage	No	Yes	Yes	Yes
Aneuploidy	No	Yes	Yes	Yes

current criteria. This is particularly the case for the *in vitro* mammalian assays. For the cytogenetic assays, it was once the practice to use concentrations up to the level where the cells no longer could divide and there were no metaphases to score. Current practice is to use cytotoxicity levels up to approximately 50%. In the early studies, there was often no measure of cytotoxicity. This is particularly true of the data obtained by the National Toxicology Program (NTP). A very large number of chemicals are included in the NTP database for cytogenetic endpoints and because there is no cytotoxicity measure, one cannot assess whether chemicals would be positive or negative using current protocols. For the MLA, it is important that experiments are conducted using optimal growth conditions to recover all of the mutants. Much, but not all of the early data, for instance from the NTP contracts, was obtained using conditions that we now know did not provide for optimal mutant recovery. Fortunately, the data collected for the NTP MLA assays include all of the observations required to reevaluate these data using current standards.

Once one assesses the quality of the data and assembles the available and acceptable data, the evaluation can begin. Ideally, there will be a complete data set for the standard assays and also information from other perhaps less widely used assays. If this is the case, there should be ample information to assess whether the chemical can cause any type of genetic damage. Unfortunately, unless the evaluation was done in a systematic manner, specifically for use in hazard identification, all of the important information may not be available. It is critical to remember that lack of information does not indicate lack of activity. That is, if there are little or no data, it is not correct to conclude that the chemical is not a mutagen or to say that there is no evidence that it is a mutagen. While the latter statement is technically correct, it should be clarified with a statement to indicate that there is no information upon which to judge if the chemical is a mutagen.

When there is information from multiple endpoints and multiple assays, the strategy for conducting a weight-of-the-evidence assessment will depend upon the question that is being addressed. It is, of course, much more difficult to draw conclusions as to whether a chemical induces cancer through a mutagenic mode-of-action than to determine whether the chemical is mutagenic in the assays that have been used.

For a cancer mode-of-action assessment of a chemical known to be a carcinogen, the key piece of information is whether the induced tumors result from mutations induced by the chemical. While there is substantial progress developing techniques that can quantify specific oncogene or tumor suppressor gene mutations (see Section "Cancer Biomarkers"), it is generally not possible to adequately address this key question. In the absence of this information, one can use the available data to draw some conclusions and generate hypotheses. The next best evidence is to assess the ability of the chemical to induce mutations *in vivo* in the tumor target tissue. This can be done using transgenic mutation assays—primarily the Big Blue, MutaMouse and *gpt* delta assays. It is, however, important to recognize that there are currently no techniques that can be used to completely define the lack of mutagenicity of a chemical *in vivo* in either the tumor target or nontarget tissue. Furthermore, at this point, the techniques that can be used *in vivo* in target tissues do not cover the full range of possible mutations important in the etiology of cancer. While the *gpt* delta assay detects the broadest range of genetic events, it does not detect large-scale deletions and chromosome rearrangements including mitotic recombination. It is important to determine if the chemical can induce both point mutations and chromosomal events, and the only possible way to accomplish this is to use both a gene mutation assay and a cytogenetic assay. While, as indicated earlier, the transgenic mutation models can be used to assess damage in both target and nontarget tissues, the cytogenetic assays are primarily limited to blood cells, which may not be the target tissue.

While in general one should place more weight on positive results found *in vivo*, the *in vitro* data are still critical in the overall assessment. Chemicals that are not positive *in vivo*, yet clearly demonstrate genotoxicity *in vitro* should not automatically be ruled as nonmutagenic. As far as the weight that should be placed on the various endpoints, mutation data should be considered to be more relevant than other endpoints; that is, mutation data should receive more weight than primary DNA interaction data. One exception to this would be when the primary DNA interaction data are obtained *in vivo* in the target organ. Positive data in the target tissue provide strong evidence that the chemical in question can reach the target tissue and can interact with the DNA.

Mode-of-Action for Cancer

Historically, cancer has been thought to result from the induction of mutations which occur with linear dose–response kinetics. Therefore, there would be no "safe" dose of a carcinogen. When estimating risk from carcinogens, it had been the norm to apply a linear extrapolation model to the available cancer data. Generally, the available data are from rodent bioassays, although there are situations where human data are available. In particular, the risk assessment for arsenic in drinking water was based on human data from a study conducted in Taiwan.

In 2005, the U.S. EPA issued a revision of its Cancer Risk Assessment Guideline, which contains two significant changes from previous cancer risk assessment practice. Firstly, it provides for utilizing information as to how the chemical actually causes tumors and then selecting an extrapolation model based on this mode-of-action. It also recognizes the possibility that children may be more susceptible to the effects of mutagenic chemicals. Therefore, it may be necessary to consider the cancer risk might be greater at a particular

exposure level in children than it would be in adults, particularly if the mode-of-action includes the induction of mutation. These advancements in cancer risk assessment mean that it becomes important to understand how a carcinogen actually induces tumors, in particular whether a carcinogen has a mutagenic or a nonmutagenic mode-of-action.

This new view of cancer and cancer risk assessment requires new thinking and a new way of structuring the analysis of available data and in determining what data might be missing for a complete evaluation. The concepts of "key events" and "key event frameworks" provide a means for assembling the data and understanding the biology underlying the induction of tumors.

The induction of tumors requires two basic key events: an increase in the frequency of cancer-causing mutations and cell proliferation. The timing of these two events is critical in establishing the appropriate mode-of-action. If the chemical directly causes a mutation as the first step (key event) in the biological effect cascade, then the chemical is a "mutagenic" carcinogen. On the other hand, if the chemical first causes some other biological effect, such as mitogenesis, hormone disruption, epigenetic alterations of cell cycle control, or binding to cellular receptors that results in disruptions to signal transduction, and that ultimately causes an increase in cell proliferation, then the mode-of-action is nonmutagenic. With a nonmutagenic mode-of-action, the frequency of cancer mutations may be increased by the proliferation of cells, thereby increasing the chance that preexisting or spontaneous mutation may contribute to the initiation of the cancer process. Cells containing cancer-causing mutations often have a proliferative advantage over nonmutant cells, and that clonal expansion of spontaneous mutations can result in rapid cell division and therefore relatively "high" mutant frequencies in the target tissues.

In the 2005 U.S. EPA Cancer Guidelines, the concept of using a modification of the Hill Criteria (from the field of epidemiology) is outlined. This provides for an evaluation of the data with regard to its strength, consistency and specificity of association, dose–response concordance, temporal relationships, and biological plausibility and coherence. While much of the data for establishing mode-of-action (e.g., measurements of apoptosis and cell proliferation) are obtained using rodent tumor target tissues(s), most of the available data to assess whether a chemical is a mutagen come from a number of genetic toxicology test systems, many of which are *in vitro*, and which were specifically designed for hazard identification. It is thus necessary to conduct a weight-of-the-evidence evaluation of all of the genetic toxicology data to draw conclusions. Unfortunately, many such evaluations fail to take all of the assay properties into account and more often than not are a simple assessment of the number of positive and negative assays, with the majority call used as the final answer.

The general science of mode-of-action determination is evolving rapidly. To obtain optimal information, it is important that any new mode-of-action experiments use an experimental design and *in vivo* exposure that directly addresses the mode-of-action assessment rather than simple hazard identification. The major distinction between the two experimental designs is that while hazard identification generally uses maximum tolerated doses and a variety of tissues, often focusing on the liver, peripheral blood or bone marrow, a mode-of-action study should be designed based on the approach used in the positive cancer bioassay; that is, the doses and the tissues evaluated should be selected based on those used in the cancer bioassay and in the tissues showing tumors. The route of exposure for these studies should mirror that used in the cancer bioassay and the length of treatment should be such that it is possible to assure that if mutations are induced, they reach sufficient numbers to be enumerated. In this context, it is important to remember that many/most chemicals require long exposures to the rodents used in the cancer bioassays before preneoplastic lesions or tumors are induced.

The most scientifically rigorous approach for determining mode-of-action is to assess all of the possible biological effects caused by the chemical treatment and to obtain information as to the temporal sequence of events and their dose responses. It is helpful if the etiology of the tumor includes preneoplastic lesions that appear prior to the development of tumors. This allows for an experimental design using biomarkers relevant for the tumor etiology and treatments that can be less than the 2 years generally used for the cancer bioassay. While the most appropriate genetic targets for this assessment would be the oncogene(s) and/or tumor suppressor gene(s) involved in the etiology of the specific tumor, this is currently not technically possible. There are, however, new techniques for evaluating the mutant frequency for a small number of specific oncogene/tumor suppressor genes and in the future, it should be feasible to directly answer the question as to whether an induced mutation is the key event in the etiology of the tumor. Currently, the best approach is to use surrogate genes, generally the transgenic models, and to address the question as to whether the chemical in question can induce mutations in the target tissue and if there is dose response concordance between the induction of the mutations and the induction of tumors. A strategy for using *in vivo* mutation data to inform cancer mode-of-action assessment is included in the reading list.

14.8 PROMISING ASSAYS UNDER DEVELOPMENT

Cancer Biomarkers: Allele-Specific Competitive Blocker Polymerase Chain Reaction (ACB-PCR)

With the goal of detecting rare mutational events, Dr. Barbara Parsons developed the sensitive ACB-PCR method for quantifying specific oncogene and tumor suppressor gene mutations.

TABLE 14.4 Specific Oncogene or Tumor Suppressor Gene Mutations that Have Been Developed for Use with the ACB-PCR Method

Species	Gene	Codon	Mutation
Mouse	H-*ras*	61	CAA→AAA
	H-*ras*	61	CAA→CTA
	K-*ras*	12	GGT→GAT
	K-*ras*	12	GGT→GTT
	K-*ras*	12	GGT→TGT
	p53	270	CGT→TGT
Rat	H-*ras*	61	CAA→CTA
	K-*ras*	12	GGT→GAT
	p53	12	GGT→GTT
	p53	271	CGT→CAT
Human	*BRAF*	600	GTG→GAG
	KRAS	12	GGT→GAT
	KRAS	12	GGT→GTT
	PIK3CA	1047	CAT→CGT

ACB-PCR involves the use of a specific PCR primer that selectively amplifies the mutant allele in conjunction with a blocker primer that prevents the amplification of the wild-type allele. The method has a sensitivity of 10^{-5}, which means that it is possible to detect a single mutant allele in the presence of 100,000 wild-type alleles. Unlike some other methods that are only semi-quantitative, ACB-PCR provides a mutant frequency because it is conducted using a set of standards with defined mutant fractions. A specific technical method must be developed for each individual mutation. To date, Dr. Parsons and her collaborators have developed methods to evaluate 14 different specific mutations (Table 14.4).

Using ACB-PCR, it has been possible to determine that tumors often carry substantial subpopulations of cells with particular mutations; that is, tumors that have been sequenced and declared to not include a specific mutation may, in fact, have that mutation in a relatively large proportion of its cells. The ability to quantify subpopulations of cells carrying particular oncogene/tumor suppressor gene mutations also provides the opportunity to use this methodology for personalized medicine. It is becoming more widely recognized that tumors can carry specific mutations that impact the choice of therapy and the overall prognosis. Currently, tumor DNA is characterized using relatively insensitive methods (like DNA sequencing), which detect mutation only when present in large proportions of tumor cells. The more sensitive ACB-PCR approach could be used to obtain quantitative data to help guide the development of personalized cancer therapies.

The ACB-PCR technology is also readily applicable to cancer mode-of-action studies. In fact, because the mutations detected by this approach can actually *be* the key mutations involved in the etiology of specific tumors, one can easily argue that this approach is superior to approaches that assess mutation in surrogate genes like *Hprt* and the transgenes (*LacI*, *cII*, and *Lac Z*). Using ACB-PCR, it has been demonstrated that the frequency of these cancer-specific mutations increase in frequency after relatively short treatment times with the carcinogens benzo(a)pyrene, aristolochic acid, and azoxymethane. Mode-of-action studies can be designed to capitalize on the capabilities of this new technology and to relate cancer-specific mutation frequencies with other events such as DNA adducts, cellular toxicity, histopathology, and changes in DNA expression. Key event frameworks can be constructed that relate these various biological effects across time and by exposure levels so that both temporality and dose–response concordance can be evaluated. Such an approach should greatly improve evaluations as to whether specific chemicals cause cancer through a mutagenic or nonmutagenic mode-of-action.

Pig-a Gene Mutation Assay

A new *in vivo* gene mutation assay using the phosphatidylinositol glycan, class A gene (*Pig-a*) is currently being developed and was the subject of a special issue of Environmental and Molecular Mutagenesis in 2011. Like the *Hprt* gene, the *Pig-a* gene is located on the X-chromosome and therefore is readily amenable for use as a mutation endpoint. There is a high degree of similarity in the gene's structure and function between species, and assays have been developed using mice, rats, nonhuman primates, and humans. This ability to evaluate the same endpoint in a number of species, including humans, makes it particularly useful for interspecies comparisons.

The product of the *Pig-a* gene is involved in the first step in the biosynthesis of glycosylphosphatidylinisotol (GPI) anchors on the cell membrane surface. While there are a number of related genes that are involved in this GPI anchor biosynthesis, the *Pig-a* gene is the only one that is located on the X-chromosome; the others are all autosomal and therefore not readily useful for detecting newly induced mutations. Also, like the *Hprt* gene, mutation in the *Pig-a* gene has a human phenotype, Paroxysmal Nocturnal Hemoglobinuria (PNH), which is an acquired genetic disorder affecting approximately 1–10 per million humans. The disease results from a somatic *PIG-A* gene mutation that occurs within a bone-marrow stem cell. These stem cells are precursors for the hematopoietic system and therefore mature blood cells, including erythrocytes, granulocytes, and monocytes, would be expected to carry the mutation once the mutated stem cell decedents differentiate.

The *Pig-a* mutation assay is based on the fact that mutant cells cannot express the GPI cell surface anchors. Because antibodies can be used to identify wild-type and mutant cells by flow cytometric methods, it is possible to directly enumerate the number of mutants without the need for mutant enrichment (mutant selection) methods. The initial reports of *Pig-a* mutation assays in rodent cells described this assay as relatively fast, inexpensive, and when

used to measure mutation in erythrocytes, requiring extremely small samples of blood. This makes it possible to measure mutation by taking small blood samples from rodents at various time points during treatment, without the need to sacrifice that animal. As is often the case in the development of new mutation assays, studies using the model alkylating agent ethyl nitrosourea (ENU) found that high mutant frequencies could be induced.

There are currently a number of laboratories conducting research to investigate various aspects of the assay including expanding the number of mutagens that are used and developing a better understanding of the assay characteristics and capabilities. There is also an international inter-laboratory trial focused on the portability of the assay, its reproducibility, and its relative sensitively to other *in vivo* genotoxicity assays. Participating laboratories first demonstrate their proficiency with the assay and to conduct experiments using ENU exposures. Laboratories then expand the number of mutagens used (including benzo(a)pyrene, *N*-methyl-*N*-nitrosourea, 4-nitroquinoline-*N*-oxide and 7,12-dimethylben(a)pyrene) and also try a 28-day experimental design, thus evaluating the potential for using the *Pig-a* assay in conjunction with standard 28-day toxicology tests.

As the *Pig-a* endpoint can be evaluated in humans, there are also studies underway to assess this capability. One study using cancer patients undergoing chemotherapy found only modest increases in *PIG-A* mutants when individuals were sampled prior to and during chemotherapy. It will be necessary to conduct additional research before final conclusions concerning the ability of this endpoint to detect induced mutations in human blood cells can be made.

14.9 SUMMARY

Mutagenesis is the study of the fundamental mechanisms by which mutations, heritable alterations in the genetic material, occur. In the context of toxicology, the study of the induction of mutations and how and which chemicals can cause mutational damage is generally referred to as genetic toxicology. This chapter was written to provide the reader with a general overview of both, the basic biology involved in the induction of mutation and how mutation is involved in human disease, thus providing a context for the importance of genetic toxicology in the general field of toxicology. It also was written to provide a historical perspective in the evolution of genetic toxicology and to provide information on the most widely used genetic toxicology tests. Specific emphasis was placed on the regulatory use of genetic toxicology data as well as some insight provided into how the assays and the interpretation of data from the assays have changed over time. The importance of, and a strategy for, conducting appropriate weight-of-the-evidence evaluations of sets of data both for hazard identification and also for mode-of-action assessments is included in the chapter.

As indicated in several sections, an OECD workgroup of international genetic toxicology experts is currently working on updating the OECD genetic toxicology testing guidelines. While it is still a work in progress, some of the revised test guidelines were provisionally approved in April 2014. The others should follow, hopefully in 2015. The expert workgroup is also drafting an overview document that will provide useful additional information beyond what is included in the individual test guidelines.

The chapter concludes with a short description of two promising new methodologies that are expected to provide new ways to expand the types of regulatory issues that can be addressed by the field of genetic toxicology.

SUGGESTED READING

Adler ID, Shelby MD, Bootman J, Favor J, Generoso W, Pacchierotti F, Shibuya T, Tanaka N. Summary report of the Working Group on Mammalian Germ Cell Tests. Mutat Res 1994;312:313–318.

Bonassi S, Ugolini D, Kirsch-Volders M, Stromberg U, Vermeulen R, Tucker JD. Human population studies with cytogenetic biomarkers: review of the literature and future prospectives. Environ Mol Mutagen 2005;45:258–270.

Bonassi S, Znaor A, Ceppi M, Lando C, Chang WP, Holland N, Kirsch-Volders M, Zeiger E, Ban S, Barale R, Bigatti MP, Bolognesi C, Cebulska-Wasilewska A, Fabianova E, Fucic A, Hagmar L, Joksic G, Martelli A, Migliore L, Mirkova E, Scarfi MR, Zijno A, Norppa H, Fenech M. An increased micronucleus frequency in peripheral blood lymphocytes predicts the risk of cancer in humans. Carcinogenesis 2007;28:625–631.

Chen T, Moore MM. Screening for chemical mutagens using the mouse lymphoma assay. In: Yan Z, Caldwell GW, editors. *Optimization in Drug Discovery: In-vitro Methods*. Totowa: Humana Press; 2004. p 337–352.

Cimino MC. Comparative overview of current international strategies and guidelines for genetic toxicology testing for regulatory purposes. Environ Mol Mutagen 2006;47:362–390.

Dearfield KL, Auletta AE, Cimino MC, Moore MM. Considerations in the U.S. Environmental Protection Agency's testing approach for mutagenicity. Mutat Res 1991;258:259–283.

Dearfield KL, Levy D, Lorge E, Moore MM, Thybaud V, Cimino MC, Custer L, Czich A, Harvey JS, Hester S, Kim JH, Kirkland D, Ouedraogo-Arras G, Schuler M, Suter W, Sweder K, Tarlo K, Van Goethem F, Witt KL. Follow-up actions from positive results on in vitro genetic testing. Environ Mol Mutagen 2011;52:177–204.

Dearfield KL, Moore MM. Use of genetic toxicology information for risk assessment. Environ Mol Mutagen 2005;46:236–245.

Dobrovolsky VN, Casciano DA, Heflich RH. *Tk*+/− mouse model for detecting *in vivo* mutation in an endogenous, autosomal gene. Mutat Res 1999;423:125–136.

Elespuru RK, Agarwal R, Atrakchi AH, Bigger CAH, Heflich RH, Jagannath DR, Levy DD, Moore MM, Ouyang Y, Robison TW, Sotomayor RE, Cimino MC, Dearfield KL. Current and future application of genetic toxicity assays: the role and value of *in vitro* mammalian assays. Toxicol Sci 2009;109:172–179.

Hayashi M, Tice RR, MacGregor JT, Anderson D, Blakey DH, Kirsh-Volders M, Oleson FB Jr, Pacchierotti F, Romagna F, Shimada H, et al. *In vivo* rodent erythrocyte micronucleus assay. Mutat Res 1994;312:293–304.

International Programme on Chemical Safety (IPCS). 2006. Transgenic animal mutagenicity assays. Available at http://www.inchem.org/documents/ehc/ehc/ehc233.pdf. Accessed July 7, 2014.

Kasper P, Uno Y, Mauthe R, Asano N, Douglas G, Mathews E, Moore MM, Mueller L, Nakajima M, Singer T, Speit G. Follow-up testing of rodent carcinogens not positive in the standard genotoxicity testing battery: IWGT workgroup report. Mutat Res 2007;627:106–116.

Kilbey BJ, Legator M, Nichols W, Ramel C. *A Handbook of Mutagenicity Test Procedures*. Amsterdam: Elsevier Scientific Publishing Company; 1977.

Lambert IB, Singer TM, Boucher SE, Douglas GR. Detailed review of transgenic rodent mutation assays. Mutat Res 2005; 590:1–280.

Li AP, Carver JH, Choy WN, Hsie AW, Gupta RS, Loveday KS, O'Neill JP, Riddle JC, Stankowski LJ, Yang LL. A guide for the performance of the Chinese hamster ovary cell/hypoxanthine-guanine phosphoribosyl transferase gene mutation assay. Mutat Res 1987;189:135–141.

Moore MM, Heflich RH, Haber L, Allen BC, Shipp AM, Kodell RL. Analysis of *in vivo* mutation data can inform cancer risk assessment. Regul Toxicol Pharmacol 2008;51:151–161.

Moore MM, Honma M, Clements J, Bolcsfoldi G, Burlinson B, Cifone M, Clarke J, Delongchamp R, Durward R, Fellows M, Gollapudi B, Hou S, Jenkinson P, Lloyd M, Majeska J, Myhr B, O'Donovan M, Omori T, Riach C, San R, Stankowski LF Jr, Thakur AK, Van Goethem F, Wakuri S, Yoshimura I. Mouse lymphoma thymidine kinase gene mutation assay: follow-up meeting of the International Workshop on Genotoxicity Testing—Aberdeen, Scotland, 2003—Assay acceptance criteria, positive controls, and data evaluation. Environ Mol Mutagen 2006;47:1–5.

Mortelmans K, Zeiger E. The Ames Salmonella/microsome mutagenicity assay. Mutat Res 2000;455:29–60.

Organisation for Economic Co-operation and Development (OECD). 1997a. OECD guideline for the testing of chemicals: *in vitro* mammalian chromosome aberration test. Available at http://www.oecd.org/chemicalsafety/testing/50108781.pdf. Accessed July 7, 2014.

Organisation for Economic Co-operation and Development (OECD). 1997b. OECD guideline for the testing of chemicals: mammalian erythrocyte micronucleus test. Available at http://www.oecd.org/chemicalsafety/risk-assessment/1948442.pdf. Accessed July 7, 2014.

Parsons BL, Marchant-Miros KE, Delongchamp RR, Verkler TL, Patterson TA, McKinzie PB, Kim LT. 2010. *Defining normal and pathological levels of tumor-associated mutations*, Regulatory Research Perspectives: Impact on Public Health, 9. Available at http://www.fda.gov/downloads/AboutFDA/CentersOffices/OC/OfficeofScientificandMedicalPrograms/NCTR/WhatWeDo/NCTRPublications/UCM217030.pdf. Accessed July 7, 2014.

Phillips DH, Farmer PB, Beland FA, Nath RG, Poirier MC, Reddy MV, Turteltaub KW. Methods of DNA adduct determination and their application to testing compounds for genotoxicity. Environ Mol Mutagen 2000;35:222–233.

Russell LB, Russell WL. Frequency and nature of specific-locus mutations induced in female mice by radiations and chemicals: a review. Mutat Res 1992;296:107–127.

Tice RR, Agurell E, Anderson D, Burlinson B, Hartmann A, Kobayashi H, Miyamae Y, Rojas E, Ryu JC, Sasaki YF. Single cell gel/comet assay: guidelines for *in vitro* and *in vivo* genetic toxicology testing. Environ Mol Mutagen 2000; 35:206–221.

United States Environmental Protection Agency (USEPA). 2005. Cancer risk assessment guidelines. Available at http://www.epa.gov/raf/publications/pdfs/CANCER_GUIDELINES_FINAL_3-25-05.PDF. Accessed July 7, 2014.

United States Food and Drug Administration (USFDA). 2000. Bacterial reverse mutation test. Available at http://www.fda.gov/Food/GuidanceRegulation/GuidanceDocumentsRegulatoryInformation/IngredientsAdditivesGRASPackaging/ucm078336.htm. Accessed July 7, 2014.

United States Food and Drug Administration (USFDA). 2001. Mouse lymphoma thymidine kinase gene mutation assay. Available at http://www.cfsanfdagov/~redbook/redivc1c.html. Accessed July 7, 2014.

United States Food and Drug Administration (USFDA). 2003. *In vitro* mammalian chromosomal aberration test. Available at http://www.cfsanfdagov/~redbook/redivc1b.html. Accessed July 7, 2014.

15

CHEMICAL CARCINOGENESIS

JAMES E. KLAUNIG

In the United States, cancer ranks as the second leading cause of death. Over 1 million new cases of cancer are diagnosed each year and more than 1½ million Americans die yearly from cancer. Most human cancers surface in the fifth, sixth, and seventh decade of life. However, cancer, usually genetically related, is also seen in children and younger adults. Over 50% of American men and approximately one-third of American women will develop cancer during their lifetimes. The risk of developing most types of cancer can be reduced or at least delayed by lifestyle changes including limiting exposure to the sun, cessation of smoking, eating a healthy diet, and becoming more active physically. Besides lifestyle, other causes of cancer have been established including exposure to infectious agents, radiation, and xenobiotic chemicals. Estimates suggest that 70–90% of all human cancers have a linkage to environmental, dietary, and behavioral factors (Table 15.1). Cancer is a disease of cellular mutation and aberrant cell growth. Exposure to a chemical or physical agent is usually at the root of the cancer induction whether the cause of cancer is the result of lifestyle, occupational, pharmaceuticals, or environmental factors. While our understanding of the mechanisms by which a normal cell changes to a malignant cell has progressed considerably in the last four decades, we are still naïve to many of the steps involved in the cancer process and to the means to prevent and cure this disease.

This chapter will address:

- History of cancer research
- Nomenclature and definitions of neoplasia
- Classification of carcinogens
- Multistage carcinogenesis
- Proto-oncogenes and tumor-suppressor genes
- Polymorphisms in carcinogenesis
- Causes of cancer
- Inorganic carcinogens
- Nongenotoxic (epigenetic) carcinogens
- Causes of human cancer
- Test systems for carcinogenicity assessment
- Classification evaluation of carcinogenicity in humans
- Properties of carcinogenic chemicals

15.1 BACKGROUND

Historically, there has been an association of epidemiology between the induction of cancer in human and chemical exposure (Table 15.2). For example, Percival Pott noted in the latter part of the eighteenth century a cause-and-effect linkage between scrotal cancer and an occupation—chimney sweeps. Pott suggested that the daily exposure to chimney soot was directly related to the cancer induction in these workers. A 100 years later, Butlin following up on Potts work found that European chimney sweeps had a much lower incidence of scrotal cancer. He hypothesized that the difference was due to age of the workers at the start of the exposure (young boys in England, adults in Europe) as well as better hygienic habits of the European sweeps. These findings illustrate two important foundations of chemical carcinogenesis that relate to cancer induction and incidence. One factor is the relationship between age of exposure and cancer induction, whereby young subjects are more sensitive to the carcinogenic effects of a chemical. A second factor is

Principles of Toxicology: Environmental and Industrial Applications, Third Edition. Edited by Stephen M. Roberts, Robert C. James and Phillip L. Williams.

the duration of exposure, where exposure to the British sweeps began in their youth and thus lasted longer in duration than the European sweeps initially exposed as adults. Rehn, in another occupational study in the late 1800s showed a linkage between the manufacturer of aniline dyes used for dyeing fabric and the induction of bladder cancer in dye workers. Subsequent work in the twentieth century showed that aniline dyes contained aromatic amine compounds (2-napthylamine and benzidine) that were responsible for the association between bladder cancer in humans and exposure to aniline dyes and helped to define the specific chemicals in the dyes responsible for the cancer induction. These chemicals were later found to react with nuclear DNA and to produce mutations in the DNA following metabolite modification. The observed epidemiological studies on cancer induction in humans following exposure to chemicals and radiation gave rise to laboratory studies, frequently using rodents as the model to further understand the cause and pathological changes seen in the cancer process. The development of these rodent models was paramount in the development of the multistage cancer models used today and to the discovery of therapeutic and prevention modalities for humans. The rodent cancer models have undergone constant refinement and advancement, recently in incorporating current molecular and cellular aspects of cancer formation. Initial studies by Yamagiwa and coworkers in Japan showed an association between coal tar and skin cancer using rodents. Kennaway and Hieger extracted specific chemical carcinogens from the coal tar that were responsible for the cancer induction in the rodent skin. These studies supported the earlier epidemiological observations of Pott and Butlin linking human exposure to soot and cancer. Similarly, Hueper and colleagues showed that the specific chemical components of the aniline dyes were responsible for the induction of bladder cancer in humans using a canine model. These findings correlated with the epidemiological studies by Rehn on the aniline dye workers. Berenblum and coworkers showed in a mouse skin model that the development of cancer required multiple steps and specific chemical agents could function at each of these steps; initiation and promotion. The Millers, in the 1950s showed a relationship between chemicals and their ability to bind to macromolecules (including DNA) with their carcinogenic ability. Thus these studies, using epidemiological and occupational approaches coupled with investigations in experimental laboratory models, have confirmed a linkage

TABLE 15.1 Human Cancer Death Rates Linked to Various Exposure Factors

Exposure	Percent Human Cancer Deaths
Infection	10
Occupational	4
Medicines and medical procedures	1
Sexual behavior	7
Food additives	1
Diet	35
Alcohol	4
Tobacco	33
Pollution	2
Geophysical factors	3

TABLE 15.2 Selected Historical Events in Chemical and Physical Carcinogenesis

Date	Investigator(s)	Finding
1775	Pott	Reported on a linkage between English chimney sweeps and scrotal cancer
1875	Thiersch	Noted a linkage between sunlight exposure and skin cancer
1879	Harting and Hesse	They reported on an association between uranium miners and the development of lung cancer
1892	Butlin	Studies with European chimney sweeps showed a distinct difference in cancer induction compared to English sweeps; he attribute this to age of workers and hygiene
1895	Rehn	Noted an increased incidence of bladder cancer in workers involved in the manufacture of aniline dyes
1902	Frieben	Proposed that exposure to X-rays was linked to cancer in humans
1915	Davis	Proposed that there was a linkage between oral cancer in pipe smokers and betel nut chewers
1928	Yamagiwa, Ichikawa, and Tsusui	Showed in an experimental rodent model, the induction of skin tumors by coal tar
1930	Kennaway and Hieger	Showed that the tumor induction of skin cancer in rodents by coal tar was the result of a specific chemical, dibenz[a,h] anthracene
1934	Wood and Gloyne	Proposed a linkage between the metals arsenicals, beryllium, and asbestos and human cancer
1938	Hueper, Wiley, and Wolfe	Extended the work of Rehn, showing that the induction of urinary cancer by aniline dyes was due to specific chemicals, in this case 2-naphthylamine
1941	Berenblum, Rous, MacKenzie, and Kidd	Showed experimentally, the two stage (initiation and promotion) model of cancer induction using the mouse skin
1951	Miller and Miller	Showed that the effects of chemical carcinogens was due to the carcinogen binding of the chemical or its metabolite to cellular macromolecules

between exposure to carcinogenic chemicals and physical agents and human cancer induction.

15.2 NOMENCLATURE AND DEFINITIONS OF NEOPLASIA

An understanding of the cancer process requires an understanding of the scientific terms involved in defining neoplasia (Tables 15.3 and 15.4). The terms cancer, tumor and neoplasia are often used interchangeably but in proper terminology neoplasia refers to new growth or autonomous new growth of a tissue (Table 15.3). A neoplastic lesion is referred to as a neoplasm. Neoplasms can be either benign or malignant in behavior. Benign neoplasms are lesions characterized by expansive growth, frequently exhibiting slow rates of proliferation that do not invade surrounding tissue or other organs. In contrast, a malignant neoplasm (cancer) demonstrates invasive growth characteristics, capable of spreading throughout the body. Metastases are growths in other tissues and organs that come from the cells of the primary neoplasm. A *carcinogen* is an agent, chemical or physical, which causes or induces neoplasia. This definition has been expanded to include an agent whose administration to previously untreated animals leads to a statistically significant increased incidence of neoplasia of one or more histogenetic types as compared with the incidence of the appropriate untreated control animals. Thus, both benign and malignant neoplasms are included in this definition of carcinogens. The term tumor describes a lesion that may or may not be neoplastic, and is characterized by swelling or an increase in size. In its truest sense, the term *cancer* describes those neoplasms that are malignant.

TABLE 15.3 Definitions of Neoplasia

Neoplasia	New growth or autonomous growth of tissue
Neoplasm	The lesion resulting from the neoplasia
Benign	Lesions characterized by expansive growth, frequently exhibiting slow rates of proliferation that do not invade surrounding tissues
Malignant	Lesions demonstrating invasive growth, capable of metastases to other tissues and organs
Metastases	Secondary growths derived from a primary malignant neoplasm
Tumor	Lesion characterized by swelling or increase in size, may or may not be neoplastic
Cancer	Malignant neoplasm
Carcinogen	A physical or chemical agent that causes or induces neoplasia
Genotoxic	Carcinogens that interact with DNA resulting in mutation
Nongenotoxic	Carcinogens that modify gene expression but do not damage DNA

TABLE 15.4 Nomenclature

Tissue of Origin	Benign Neoplasm	Malignant Neoplasm
Epithelial Tissue		
Skin	Squamous cell papilloma	Squamous cell or carcinoma
Lung	Bronchial adenoma	Bronchogenic carcinoma
Kidney	Renal tubular adenoma	Renal cell carcinoma
Liver	Liver cell adenoma	Hepatocellular carcinoma
Bladder	Transitional cell papilloma	Transitional cell carcinoma
Testis		Seminoma
Lymph nodes		Lymphomas
Mesenchymal Tissue		
Smooth muscle	Leiomyoma	Leiomyosarcoma
Striated muscle	Rhabdomyoma	Rhabdomyosarcoma
Connective tissue	Fibroma	Fibrosarcoma
Endothelial cells	Hemangioma	Hemangiosarcoma
Bone	Osteoma	Osteosarcoma
Mesothelium		Mesothelioma
Melanocytes		Malignant melanoma

For classifying neoplasms, both, the tissue of origin and the characteristics of the type of tissue are incorporated into the nomenclature (Table 15.4). For benign neoplasms, the tissue in which the lesion is developed is frequently followed by the suffix "oma." For example, a benign neoplasm from fibroblasts would be termed *fibroma*, and a benign neoplasm from the glandular epithelium would be termed an *adenoma*. Malignant neoplasms from epithelial origin are called *carcinomas*, while those derived from mesenchymal origin are referred to as *sarcoma*. Thus, a malignant neoplasm of fibrous tissue would be a *fibrosarcoma*, while that derived from bone would be an *osteosarcoma*. Similarly, a malignant neoplasm from the liver would be a *hepatocellular carcinoma*, while that derived from skin, referred to as a *squamous cell carcinoma*. Preneoplastic lesions have also been observed in a number of target organs in both animal models and humans and reflect an early reversible lesion in neoplasm progression. The characterization and study of preneoplastic cells has led to further understanding of the relationship process of cancer formation.

15.3 CLASSFICATION OF CARCINOGENS

Agents that cause cancer (carcinogens) can be chemicals, viruses, hormones, radiation, or solid-state materials. To be considered a carcinogen, the agent needs to either produce new neoplastic growth in a tissue or organ or increase the incidence and/or multiplicity of background spontaneous neoplastic formation in the target tissue.

Carcinogens have been divided into two general categories based on the mechanism of action by which

TABLE 15.5 Characteristics of Genotoxic and Nongenotoxic Carcinogens

Genotoxic carcinogens
DNA reactive (directly or indirectly (requiring metabolism))
Mutagenic
Can be complete carcinogens
Tumorigenicity is dose-responsive
For regulatory purposes no threshold is defined
Can be complete carcinogens
Function at initiation and progression stage of cancer process
Nongenotoxic carcinogens
Non-DNA reactive
Nonmutagenic
Modulates gene expression
A dose-responsive threshold is seen
Reversible (dose- and duration-dependent)
Tumorigenicity is dose-responsive
May function at tumor promotion stage
Species, strain, tissue specificity in tumor formation (dependent on metabolism)

carcinogen functions. These have been labeled as genotoxic and nongenotoxic (or epigenetic). Genotoxic carcinogens, as the term implies, interact physically with DNA to damage or change its structure resulting in mutational event. Nongenotoxic (epigenetic) carcinogens may effect DNA expression without modifying or directly damaging DNA structure, or may create a situation in a cell or tissue that makes it more susceptible to DNA damage from other sources. Common features of genotoxic and nongenotoxic carcinogens are shown in Table 15.5. Classification of carcinogens based on mechanism has been important in developing meaningful scientifically based approaches to determining the relative human cancer risk after exposure to the cancer-causing agent.

The majority of genotoxic, DNA reactive chemical carcinogens are found as parent compounds or procarcinogens. Procarcinogens are relatively stable chemicals that require subsequent metabolism to be carcinogenic. Metabolism occurs either in the liver or in the target tissue (where the cancer arises) itself. Since they require metabolic activation, they are referred to as indirect-acting carcinogens. A direct-acting carcinogen does not require metabolic activation. Historically, the terms procarcinogen, proximate carcinogen, and ultimate carcinogen refer to the parent compound (procarcinogen) and its metabolite forms, either intermediate (proximate carcinogen) or final (ultimate carcinogen) that reacts with DNA (Figure 15.1). The ultimate form of the carcinogen is the chemical species that produces sufficient DNA damage to induce mutation. The ultimate carcinogenic chemical forms of the most highly studied genotoxic carcinogens have been identified. Indirect-acting genotoxic carcinogens usually produce their neoplastic effects, not at the site of exposure (as seen with direct-acting genotoxic carcinogens) but at the target tissue where the metabolic activation of the chemical occurs. Examples of selected direct and indirect genotoxic carcinogens showing their pro and ultimate forms are shown in Table 15.6.

The reaction of a carcinogen with genomic DNA either directly or indirectly may result in DNA adduct formation or direct DNA damage, and produces a mutation if not repaired (see Chapter 14). Many ultimate carcinogens that produce mutation are alkylating electrophiles. Several mechanisms of mutagenesis from exposure to genotoxic carcinogens are known to occur. These include transitions and transversions, frame shift mutations, and DNA strand breaks as already discussed in Chapter 14.

Transitions are a substitution of one pyrimidine by the other or one purine by the other, while a transversion occurs when a purine is replaced by a pyrimidine or a pyrimidine is replaced by a purine. Carcinogens can induce transitions and transversions when carcinogen adducts are formed and are misread during the DNA repair process. A shift, (frame shift mutation) may also result when the carcinogen DNA adduct is formed on a nucleotide. In a third case, DNA strand breaks can occur as a result of excision repair mechanisms that are incomplete during DNA replication. All of these events are dependent upon the location of the adducts in the genomic DNA, when in the cell cycle the adducts are formed, and the type of DNA repair enzymes that respond to the damage. While carcinogen DNA adducts can be formed at all sites in DNA, the most common sites of alkylation include the N7 of Guanine, the N3 of adenine, the N1 of adenine, the N3 of Guanine, and the O6 of guanine (Figure 15.2).

Besides alkylation, another common modification to DNA is the hydroxylation of DNA bases. Oxidative DNA adducts have been identified in all four DNA bases (Figure 15.3), with 8-hydroxyguanine being the most prevalent oxidative DNA adduct formed and measured. The source of oxidative DNA damage is typically free radical reactions that occur endogenously in the cell or from exogenous sources. DNA damage and mutation induced through oxidative DNA adducts may be responsible for the "spontaneous" formation of initiated cells and tumors.

Finally, modification of the methylation status of DNA has been shown to result in heritable changes in gene expression. Hypomethylation has been associated with increased transcription of genes, while hyper methylation produces a reduction of gene expression. Chemical carcinogens may inhibit DNA methylation by several mechanisms including single-strand breaks in the DNA, forming covalent adducts, alteration of methionine pools, and inactivation of the DNA methyltransferase.

Although a large number of adducts can be formed following exposure to chemicals, whether a particular

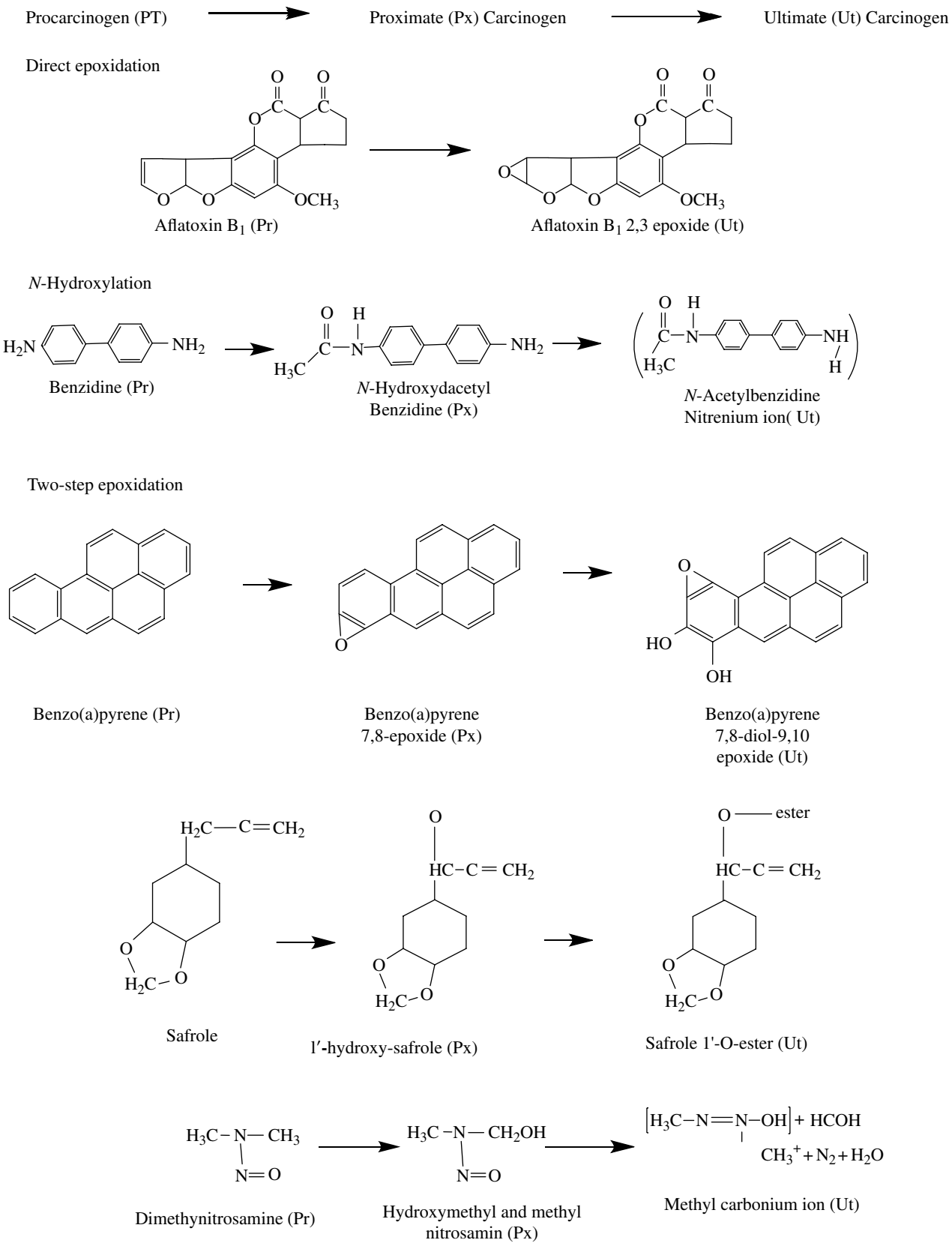


FIGURE 15.1 Structures of indirect-acting carcinogens and their DNA reactive metabolites. The procarcinogen form (parent compound) and metabolites, the proximate (Px) and ultimate (Ut) carcinogenic forms are shown.

DNA adduct will result in mutation and participate in the carcinogenesis process is dependent in part on the persistence of the DNA adduct through the process of DNA replication, which is also in part dependent upon DNA repair.

TABLE 15.6 Characteristics of DNA Reactive/Genotoxic Carcinogens

A. Direct-acting carcinogens (no metabolism needed)
Nitrogen or sulfur mustards
Methyl methane sulfonate
B-Propiolactone
1,2,3,4-Diepoxybutane
Bis-(Chloromethyl) ether
B. Indirect-acting carcinogens (require metabolic activation)
Polycyclic aromatic hydrocarbons
Aromatic amines
Nitrosoamines
Azo dyes
Hydrazines
Cycasin
Safrole
Aflatoxin B1

15.4 MULTISTAGE CARCINOGENESIS

The process of carcinogenesis involves genomic mutation in a cell with subsequent cell proliferation of the mutated cell. This multistage, multistep process has been extensively study in animal models and similar steps have been shown to occur in human cancer formation. The multiple cancer process can be exemplified using both mechanisms, using the corresponding pathology to illustrate the stages. Studies with rodent as well as the pathology of human cancers have shown that the carcinogenesis process involves a series of definable and reproducible stages. Using the most basic definition, three stages, initiation, promotion, and progression, have been shown (Figure 15.4). These stages follow a temporal sequence beginning with a normal cell and completion with a neoplastic cell. The biological characteristics of these stages are shown in Table 15.7.

Tumor Initiation

This initial step of the process involves the change of a normal cell to a mutated cell through a stable, heritable genetic change and is termed initiation. The initiation

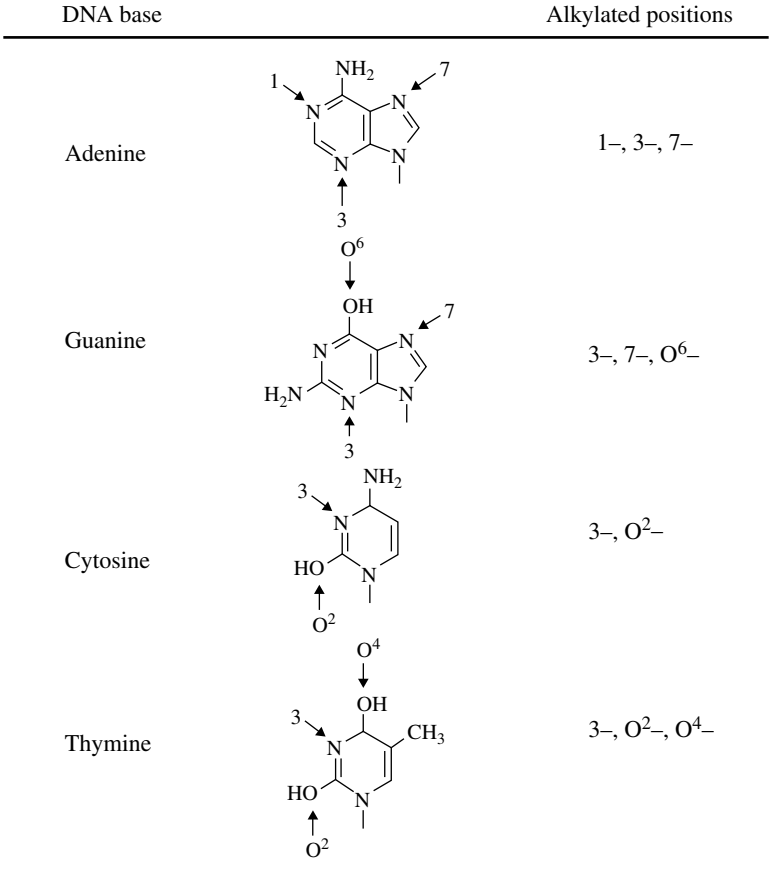


FIGURE 15.2 DNA targets for possible adduct formation by alkylating carcinogens.

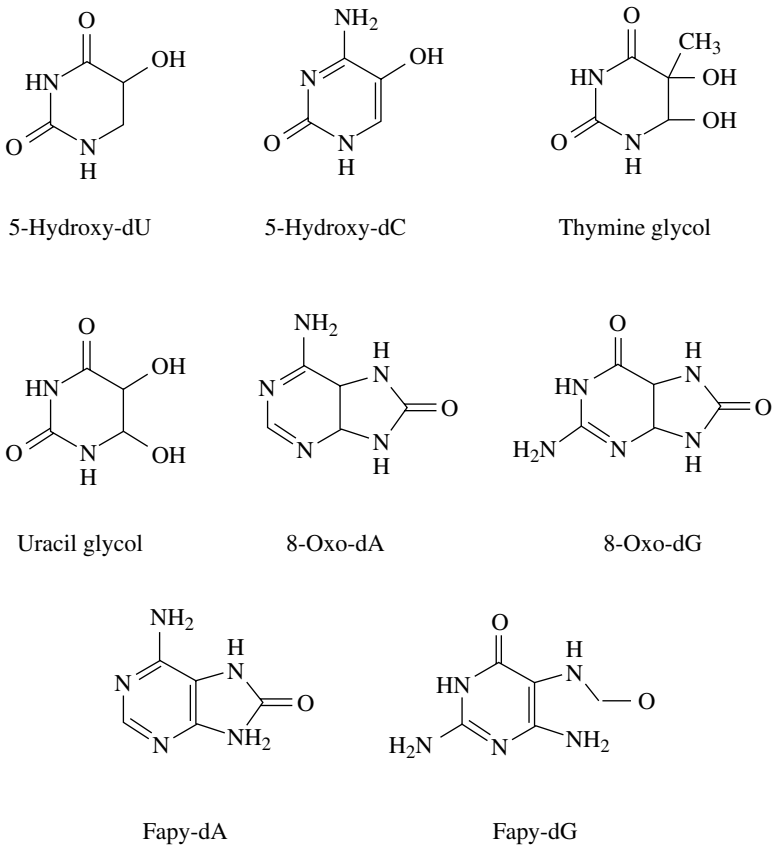


FIGURE 15.3 Structures of selected oxidative DNA bases.

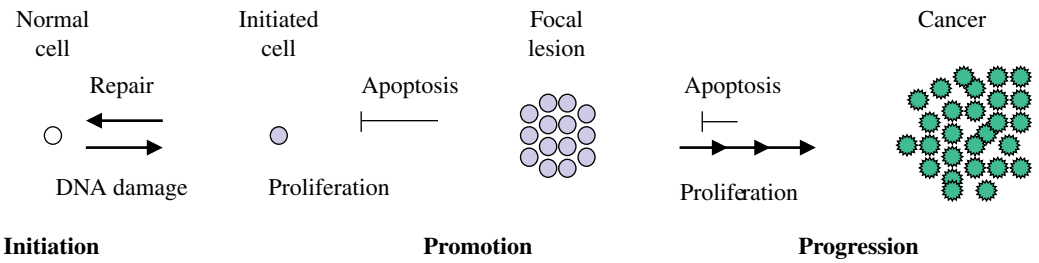


FIGURE 15.4 Multistage model of carcinogenesis involving initiation, promotion, and progression stages.

(mutation) becomes "fixed" (heritable) when the DNA adducts or DNA damage are not repaired completely prior to DNA synthesis. Once formed, initiated cells can remain in a nondividing state, or the mutations may be incompatible with viability of the cell and the cell dies, or the cell may proceed through additional cell divisions, resulting in the proliferation of the initiated cell. Chemical carcinogens that covalently bind to nuclear DNA and form adducts that result in mutations are initiating agents. Examples of initiating carcinogens include compounds such as polycyclic hydrocarbons and nitrosamines, biological agents, certain viruses, and physical agents such as X-rays and UV light. Many of the chemical carcinogens that function as initiators are indirect-acting genotoxic compounds requiring metabolic activation by the target cell to produce an ultimate form of the carcinogen that is then able to bind to nuclear DNA, resulting in the formation of a DNA adduct.

Many compounds that work at the initiation stage have the ability to function at all stages of the cancer process and are thus complete carcinogens. It is important to note that the stage of initiation, that is, the formation of a mutated cell, can also occur spontaneously through misrepair of normally acquired DNA damage during DNA replication. The cell formed during this spontaneous initiation process might have a similar alteration of its DNA sequence to that observed with physical and chemical agents. Agents that work at the tumor initiation stage are

TABLE 15.7 Characteristics of Multistep Process of Carcinogenesis

Initiation
DNA interaction
Involves a mutation
DNA reactive/genotoxic
One cell division necessary to lock-in mutation
Modification is not enough to produce cancer
Not reversible
Single dose of compound can produce a mutation
Promotion
No direct DNA modification
Nongenotoxic
No direct mutation (although repeated induction of mitosis may result in misrepair and production of new mutations)
Multiple cell divisions necessary
Clonal expansion of the initiated cell population
Increase in cell proliferation or decrease in cell death (apoptosis)
Reversible
Multiple treatments (prolonged treatment) necessary
Dose–response threshold seen
Progression
DNA modification
Genotoxic event
Mutation, chromosome disarrangement
Changes from preneoplasia to benign/malignant neoplasia
Irreversible
Number of treatments needed with compound unknown (may require only single treatment)

usually genotoxic in behavior. As such their target tissue is dependent on compound metabolism and site of treatment. Frequently, tumor initiators induce tumors at multiple tissue and organ sites.

Tumor Promotion

A second step of the cancer process involves the selective clonal expansion of initiated cells that results in a preneoplastic lesion. This stage is often referred to as tumor promotion. Exogenous and endogenous agents that act on this stage are referred to as tumor promoters. Tumor promoters are not mutagenic and generally do not induce tumors by themselves. In general, they act through a modulation of gene expression that increases the number of cells, and these cells have been initiated either directly (via a mutagen) or indirectly (via an endogenous rate of misrepair) forming a preneoplastic lesion in the target organ. The tumor promotion stage of the cancer process requires continuous exposure to the tumor promoter. While initial exposure to a tumor-promoting agent produces a transient increase in cell growth, only the initiated cells within the tissue sustain this increase in cell proliferation and/or DNA synthesis with continual treatment to the agent. Repeated applications of the tumor-promoting compound induce a continual clonal expansion of the initiated cell population (Figure 15.4). The stage of tumor promotion is reversible whereby removal of the agent results in an increase in cell death (most likely via apoptosis) in the lesion with a return back to a single or small group of initiated cells. Carcinogens that function at the tumor-promotion stage exhibit a well-defined threshold in their dose–response pattern. For each agent, a certain dose and/or frequency of application must be achieved to produce the tumor promotion. Doses below this threshold do not modify gene expression sufficiently to produce the selective clonal expansion necessary for the development of a preneoplastic lesion. Multiple chemical compounds as well as physical agents have been linked to the tumor-promotion stage of the cancer process. Tumor promoters in general show organ-specific effects, for example, a tumor promoter of the liver, such as phenobarbital, will not function as a tumor promoter in the skin or other tissues.

Tumor Progression

The final stage of the carcinogenesis process involves the formation of the benign or malignant neoplasm. This stage is frequently referred to as the progression stage since it involves the progression from preneoplasia to neoplasm. During the progression stage, additional modification of genomic DNA and even chromosomal damage such as aberrations and translocations occurs, producing cells with the neoplastic lesion that are further independent from growth control of the body. An accumulation of nonrandom chromosomal aberrations and chromosome instability are hallmarks of progression.

Mutations

As noted earlier, mutations involve changes in the arrangement of the bases that comprise a gene. Two general types of mutations can occur: hereditary (germ line) mutations and acquired mutations. Hereditary mutations are defects in the genetic code, inherited by a child from its parent. Hereditary mutations are present in zygotes and, as such, are expressed in all cells of the body including sperm or eggs. As such, a hereditary mutation can be passed from generation to generation. Approximately 5–10% of human cancers are attributed to hereditary mutations. Those mutations acquired during the lifetime of the individual, usually in somatic cells, are the most frequent mutations seen in cancer accounting for 90–95% of human cancer and are induced by chemical or physical agents. Acquired mutations are found in a single cell, or the prodigy of that cell, unlike the germ line mutations.

It is important to note that mutations can occur at multiple steps of the entire cancer process. Using the three-stage model noted earlier, mutation plays an important role in the initiation and progression stage of carcinogenesis.

Mutations in an oncogene, tumor-suppressor gene, or other genes that control the cell cycle can result in a clonal cell population with a proliferative or survival advantage. The development of a tumor requires many such events, occurring over a long period of time, and for this reason human cancer induction often takes place within the context of chronic exposure to chemical carcinogens.

DNA Repair Mechanisms

Cells contain the ability to repair many types of DNA damage. The mechanisms of DNA repair while very good are not always completely effective, and sometimes residual, unrepaired DNA damage can lead to the insertion of an incorrect base during the process of DNA replication. This ultimately leads to mutation and the synthesis of altered proteins. A variety of mechanisms have evolved to effectively repair DNA damage. The more common types of DNA repair mechanisms seen in mammals are listed in Table 15.8. Repair of DNA damage does not always occur prior to cell replication, and repair of DNA damage by some chemicals is relatively inefficient; thus, exposure to chemicals that cause DNA damage typically increases the probability of mutations.

As already stated, spontaneous mutations can occur through normal cellular DNA replication mistakes. Many spontaneous mutations are point mutations where there is a change in a single base pair in the DNA sequence. The issue for mismatch repair is determining the normal DNA and the damaged DNA strand and therefore repairing the mutated strand such that the correct base pairs are restored. The formation of apurinic sites in DNA is a fairly common occurrence and spontaneous event in mammals. Cells contain apurinic endonucleases that cut DNA near apurinic sites and then these cuts are repaired by DNA polymerases and ligases. However, if these lesions are not repaired, mutations can be produced during DNA replication.

DNA containing modified bases or adducts following carcinogen exposure are also repaired by excision repair (Figure 15.5). DNA lesions that are repaired by excision repair processes include thymine–thymine dimers (UV light) as well as modified DNA regions containing chemical adducts. Chemical–DNA adducts, when formed, produce a modification in the normal shape of DNA. The DNA repair enzymes recognize these irregularities in the shape of the double helix, and attempt to repair the lesion. The bulky adducts and thymine–thymine dimers are frequently misrepaired by the excision repair process resulting in mutations.

Double-strand breaks of the DNA can also be caused by ionizing radiation and chemicals carcinogens. A cell that has double-strand breaks can be repaired by joining the free DNA ends. These double-strand breaks are correctly repaired only when the free ends of DNA rejoin exactly. The repair of double stranded DNA is confounded by the fact that a single-stranded piece of DNA from which the base-pairing is patterned is missing with double-strand breaks. Therefore, the double-stranded DNA repair involves non-homologous DNA repair (no template to base the repair on) and thus rejoining the broken ends of the two DNA molecules results in the deletion of many base pairs, thereby increasing the risk of a mutation being formed.

TABLE 15.8 DNA Repair Pathways

1. Direct reversal of DNA damage
2. Excision repair systems
Base excision repair
Nucleotide excision repair
Mismatch repair
3. Post-replication repair (recombination repair)
4. Double-strand break repair

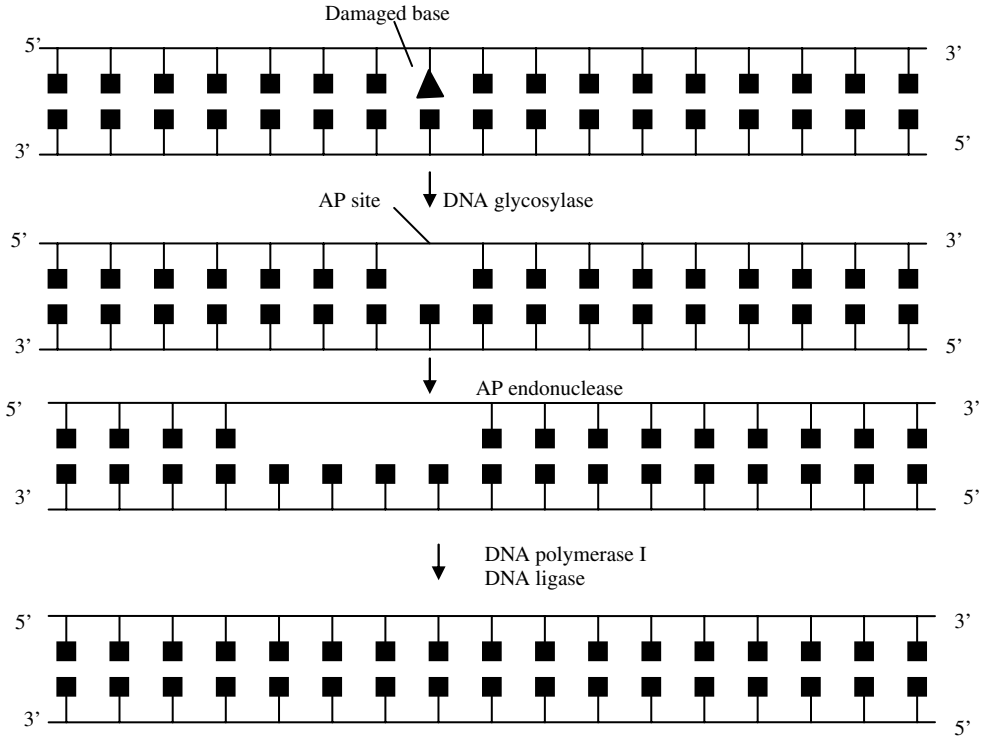
Proto-Oncogenes and Tumor-Suppressor Genes

Following the formation of a mutation, two general groups of genes have been identified as being important in the formation and progression of cancer. These include proto-oncogenes and suppressor genes. Proto-oncogenes and tumor-suppressor genes encode a number of proteins involved in cell growth differentiation, and cell death. Table 15.9 shows some of the characteristics of a number of oncogenes and tumor-suppressor genes.

Proto-Oncogenes

An oncogene is a gene encoding a protein that can transform cells in culture or produce cancer in rodents. Most of the known oncogenes actually are normal genes (hence the term proto-oncogenes) that are involved in cell signaling for maintaining cell viability and tissue homeostasis. Mutation of the proto-oncogenes results in unregulated growth control of the cell. It has been shown that a number of genotoxic, DNA reactive carcinogens are capable of producing mutations in proto-oncogenes. Several oncogene products from mutated genes have tyrosine kinase activity, which can function at growth factor receptors including EGF, PDGF, insulin, insulin-like growth factor, and FGF resulting in increased cell growth. Other oncogenes modify GTP activity (the *ras* genes—e.g., *Ha-*, *Ki-*, and *n-ras*). Mutations in R*as* affect the association of the Ras with GTP, and since Ras is active when GTP is bound, constitutive and constant Ras activation is often seen with mutated forms of R*as*. Sustained activation of Ras then leads to activation of several signal transduction proteins which in turn produces continual cell division among other effects. Other oncogene products function as nuclear transcription factors thus altering gene expression. The *jun* and *fos* family of oncogenes encode for the AP-1 transcription factor, which produces sustained

DNA repair by base excision



DNA repair by nucleotide excision

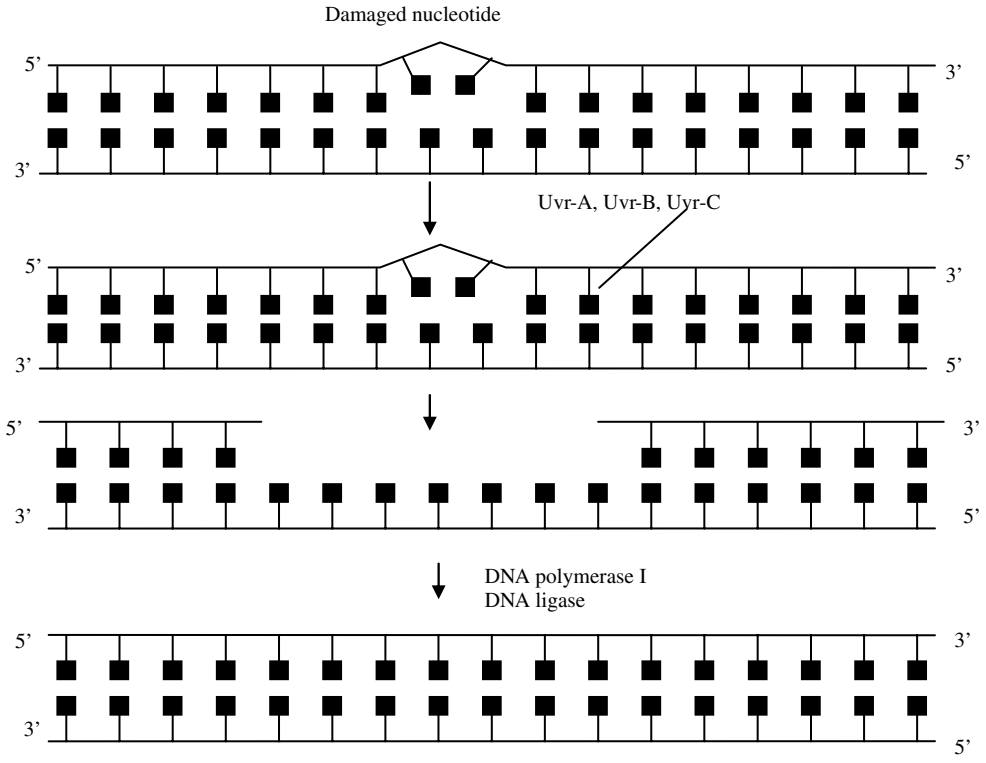


FIGURE 15.5 DNA repair.

TABLE 15.9 Common Characteristics of Proto-Oncogenes, Oncogenes, and Tumor-Suppressor Genes

Proto-Oncogenes	Oncogenes	Tumor-Suppressor Genes
Dominant trait	Dominant trait	Recessive trait
Many different tissues targeted for cancer development	Many different tissues targeted for cancer development	Displays a somewhat tissue-specificity for cancer development
Germ line inheritance rarely involved in the cancer	Germ line inheritance rarely involved in the cancer	Germ line inheritance is many times involved in the development of the cancer
Has analogy to viral oncogenes	No established analogy to oncogenic viruses	No known analogy in oncogenic viruses
Somatic mutations activated during all stages of carcinogenesis	Somatic mutations activated during all stages of carcinogenesis	Germ line mutations may be involved in initiation, but appears to be more involved in the progression stage

TABLE 15.10 Oncogenes and Tumor Site Association

Target	Oncogene	Neoplasm
Tyrosine kinases	EGFR	Squamous cell carcinoma
	PDGF	Lung carcinoma Astrocytoma
	Src	Colon carcinoma
	Neu	Breast carcinoma Neuroblastoma
	Ret	Thyroid carcinoma
	Trk	Colon carcinoma
Serine/threonine kinases	v-raf	Sarcoma
	v-mos	Sarcoma
G proteins	H-Ras	Colon and lung carcinoma
	K-Ras	Melanoma, AML, and thyroid
	N-Ras	Carcinoma and melanoma
	NF-1	Neurofibromas
Nuclear proteins	c-Myc	CML, Burkitt's lymphoma
	N-Myc	AML, breast and lung carcinoma
	L-Myc	Neuroblastoma and lung
	v-myb	Lung carcinoma
	v-jun	Sarcoma
	v-fos	Osteosarcoma

stimulation of cell growth signals. The *myc* oncogene family (*c-myc*, N-*myc*, and L-*myc*) encode for transcription factors that also are involved in cell growth. Therefore, activation of these oncogenes results in the continual cell growth of the target cell. Table 15.10 shows examples of selected oncogenes and the site of tumor formation.

Tumor-Suppressor Genes

In contrast to oncogenes, most tumor-suppressor genes code for proteins that either act as inhibitors of cell proliferation or reduce cell survival, thereby eliminating the mutated cell. One of the first suppressor genes noted was the *Rb* gene, which is seen mutated in retinoblastoma. Loss or mutational inactivation of *Rb* contributes to the development of a wide variety of human cancers.

Similarly, the p53 protein is a tumor-suppressor gene that is essential for the controlling of the cell cycle. Cells with functional p53 stop in G_1 of the cell cycle when the cell is exposed to DNA damaging agents. Cells that lack or have a mutated p53 are unable to stop the cell cycle to allow for more complete DNA repair or induce cell death. Therefore, inactivated p53 allows DNA damaged cells to continue to divide and incorporate the DNA damage in future cells created by cell division. Accumulation of p53 also leads to induction of proteins that promote programmed cell death (apoptosis) and this process prevents the proliferation of cells that are likely to accumulate multiple mutations. When p53 is mutated, this protective pathway does not function properly and the damaged DNA can replicate, producing mutations and DNA rearrangements that result in cancer. In some breast cancers, a hereditary modification of *BRCA1*, a caretaker gene that produces a protein (breast cancer type 1 susceptibility protein) which is responsible for repairing DNA. *BRCA1* is normally expressed in breast cells where it repairs damaged DNA, or induces cell death in breast cells that cannot be repaired. If *BRCA1* is modified or damaged, any breast cell damaged DNA is then not repaired properly, resulting in an increase in possible mutations of these cells and an increased risk for cancer. Mutation of *BRCA1* has been reported to increase the risk of developing breast cancer to 60% by the age of 50. Table 15.11 shows examples of selected tumor-suppressor genes and their tumor site association.

Polymorphisms in Carcinogenesis

Genetic polymorphisms arise when a gene has more than one allele. It has been found that DNA base variations occur approximately once in every 1000 base pairs, resulting in the possibility of over 1 million genetic variations between any two individuals. A single nucleotide polymorphism (SNP) is a variant in DNA that is found in more than 1% of the population. Genetic polymorphisms may be responsible for

TABLE 15.11 Examples of Tumor-Suppressor Genes and Tumor Site Association

Tumor-Suppressor Gene	Cancer
Rb1	Small-cell lung carcinoma
p53	Breast, colon, lung cancers
BRCA1	Breast carcinoma
WT-1	Lung cancer
p16	Melanoma

TABLE 15.12 Carcinogen Exposure and Cancer Risk in Humans

Susceptibility (Genotype)	Exposure to Carcinogen	Cancer Risk
↓	↓	Low risk
↓	↑	Low risk
↑	↓	Mod risk
↑	↑	High risk

the susceptibility of individuals to cancer. Of particular interest have been the SNPs of genes involved in carcinogen metabolism (activation and detoxification) and DNA repair. Polymorphisms in carcinogen activation, including P450 enzymes as well as enzymes involved in detoxification such as glutathione-S-transferases (GSTs) have been found and linked to cancer risk. For example, in nonsmokers with lung cancer, an increase in PAH-DNA adducts of polycyclic hydrocarbons (found in smoke) have been detected. The DNA adduct levels correlate with the increased expression of cytochrome P4501A1 (CYP1A1) related to an SNP for this gene. In addition, gene polymorphisms of the GSTM1 phenotype have been related to bladder, gastric, and lung cancer in humans. SNPs of DNA repair genes; in particular, oxidative DNA repair, OGG1, has been linked to increase in lung cancer in individuals with lower OGG1 expression. The importance of SNPs, cancer susceptibility, and cancer risk are illustrated in Table 15.12. If chemical carcinogen exposure is low and the genetic susceptibility for genes related to the carcinogenic susceptibility is low, then the risk for cancer will also be low. If carcinogen exposure is high and the genetic susceptibility is low, then the risk for cancer development is still likely to be low. However, if the genetic susceptibility is high and carcinogen exposure is high, then the risk for cancer formation will be increased.

15.5 CAUSES OF CANCER

As mentioned previously, there are multiple causes of cancer. These include chemicals, physical, bacterial, and viral agents from exposure via occupational, medical therapies, environmental, and lifestyle moieties. A selection of the most studied of these carcinogenic agents is discussed in the following. There exist several excellent reviews and evaluations of chemicals, lifestyle factors, and physical agents that induce cancer and their potential of human cancer risk following exposure. These include the National Toxicology Program, The International Agency for Research in Cancer (IARC) monograph series, and the CDCs ATSDR toxicological profiles.

Polyaromatic Hydrocarbons

The chemical class polyaromatic hydrocarbons (PAHs) are one of the most studied groups of carcinogens. Examples of some of the more commonly studied individual PAHs are found in charcoal broiled foods, cigarette smoke and diesel exhaust. Chemical structures of selected carcinogenic polycyclic hydrocarbons are shown in Figure 15.6. Most PAHs bind to DNA and induce carcinogenesis through genotoxic mechanisms. Benzo(a)pyrene (BaP) is the model compound of study for the PAH class of carcinogens. The metabolism of BaP that leads to tumor formation has been extensively characterized. Benzo(a)pyrene is first oxidized by cytochrome P4501A1 to form benzo(a)pyrene 7,8-oxide, which is further metabolized by epoxide hydrolase, to produce a 7,8,dihydrodiol compound. 7,8,dihydrodiol is then metabolized via further epoxidation to form 7,8-dihydrodiol-9,10-epoxide, the metabolite that is the ultimate carcinogenic chemical intermediate. This ultimate form is the one that binds to nuclear DNA. The bay region of the benzo(a)pyrene molecule is the site where most of the metabolism occurs and where the molecule forms adduct with DNA (Figure 15.7). Other PAHs beside BaP have similar structures and Bay regions and follow a similar pattern of metabolism occurs producing other PAH metabolites that bind to DNA and form adducts.

Alkylating Compounds

Alkylating agents are a large and varied class of genotoxic carcinogens that work through either direct action (no metabolism required) or indirect action (metabolism to the ultimate carcinogenic form is required). While the alkylating compounds (or their metabolites) readily react at more than 12 sites on DNA bases, the N7 position of guanine and the N3 position of adenine are the most reactive DNA sites. The nitrosamines were initially used as solvents in chemistry until their toxic effects were identified when workers developed liver damage. Studies in animal models confirmed that the nitrosamines are highly hepatotoxic and carcinogenic. Other alkylating chemicals including the nitrogen mustards have found application in cancer chemotherapy because of their high reactivity with DNA which leads to the cancer cell death. Alkylating agents produce DNA adducts and DNA strand breaks. Nitrogen mustards can produce a wide spectrum of

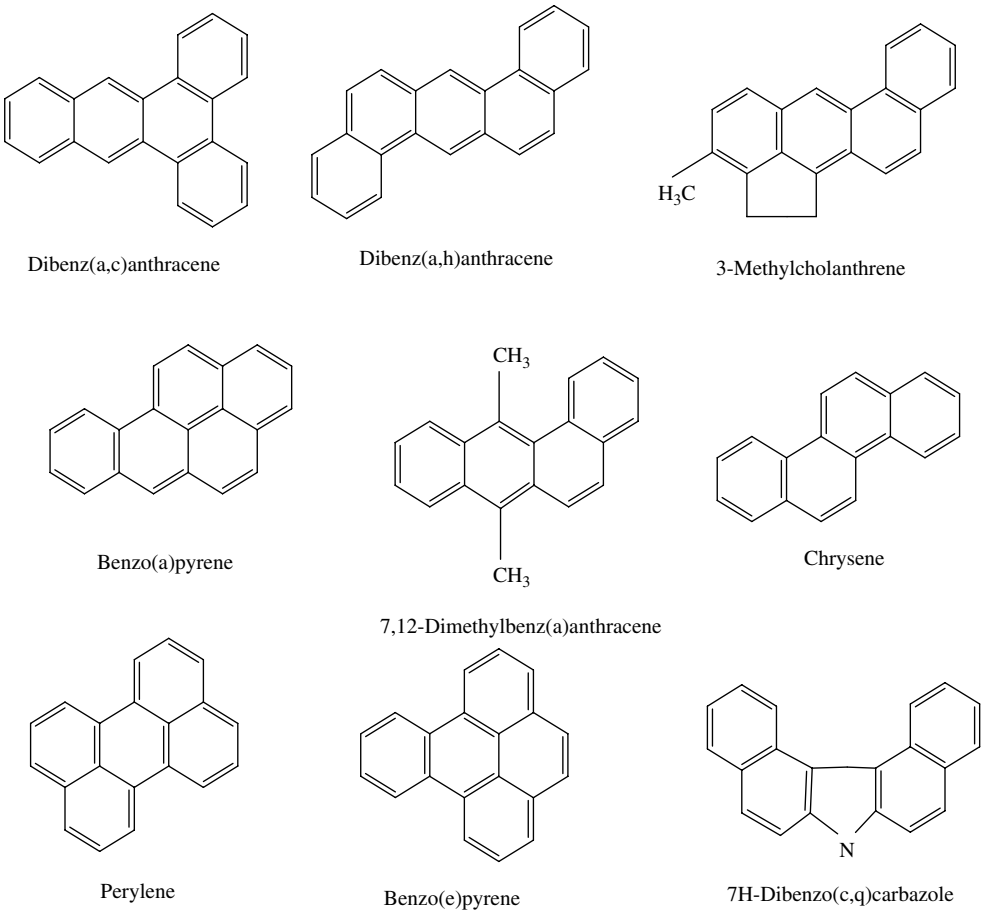


FIGURE 15.6 Chemical structures of selected carcinogenic polycyclic hydrocarbons.

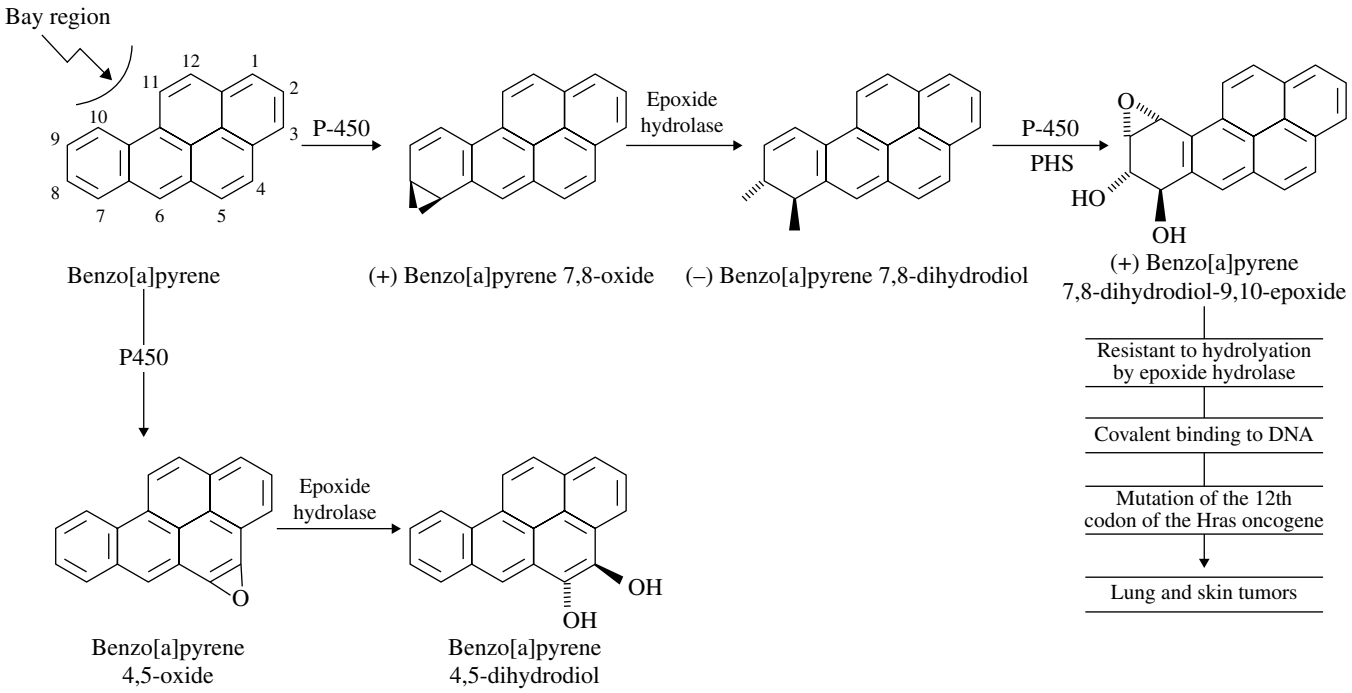


FIGURE 15.7 Metabolism of benzo[a]pyrene to the tumorigenic bay-region diol epoxide metabolite.

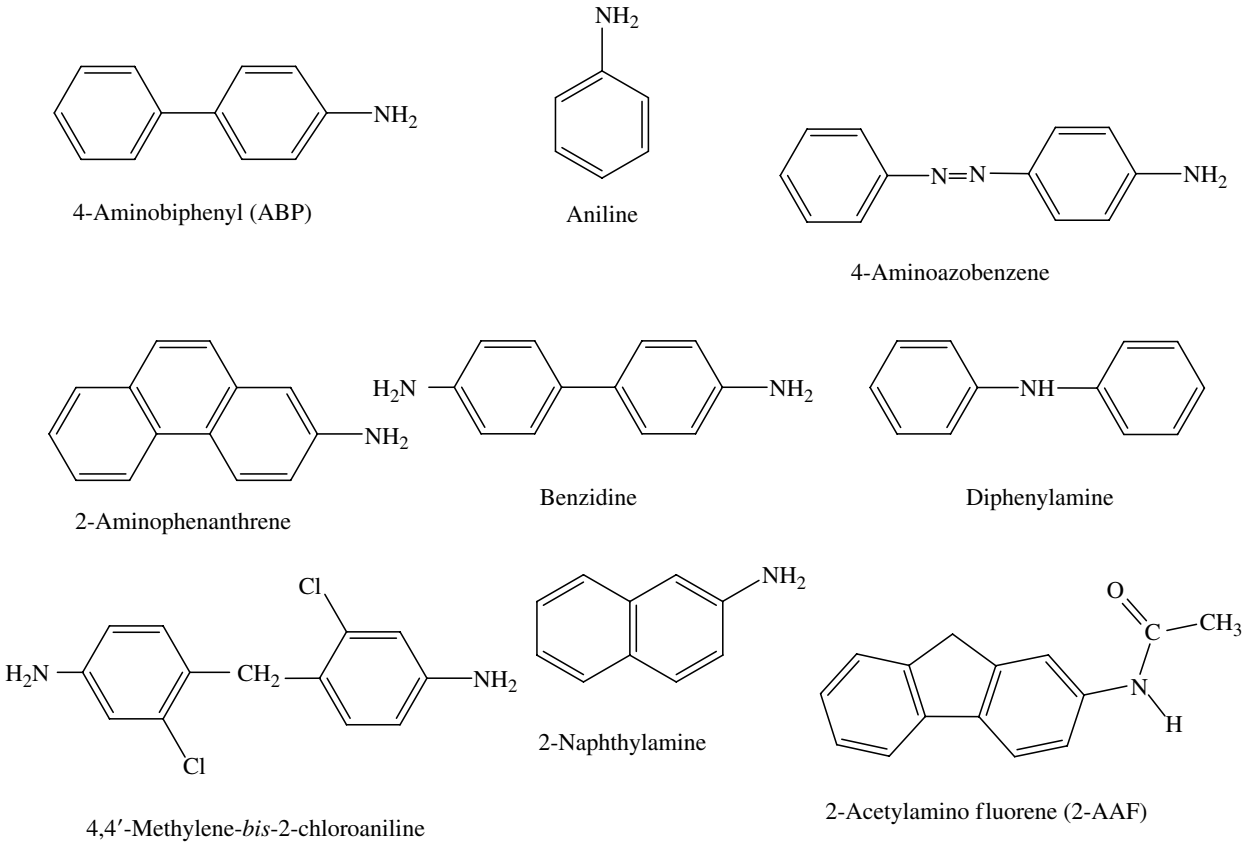


FIGURE 15.8 Selected carcinogenic aromatic amines.

mutations including base pair substitutions and deletions. Ethylene oxide and propylene oxide are also direct-acting alkylating carcinogens in rodents and are of human concern. Alkylation of DNA by ethylene and propylene oxide occurs predominantly at the N7 position of guanine. Vinyl chloride is another known rodent and human carcinogen, producing angiosarcomas in the liver in humans. Some epidemiological evidence also suggests that vinyl chloride may produce brain tumors. Vinyl chloride is metabolized by cytochrome P450 to form chloroethylene oxide, which can also form chloroacetaldehyde, both of which can alkylate DNA. Vinyl chloride and metabolites form the DNA adducts 7-(2′-oxoethyl)guanine, N2, 3-etheneoguanine, 3,N4-ethenocytosine, and 1,N6-ethenoadenine.

Aromatic Amines

Aromatic amines include a number of DNA reactive chemical carcinogens that have been used in industry as chemical intermediates and have potential human occupational carcinogenic risk. The aromatic amines undergo metabolism mainly by cytochrome P450-mediated reactions, yielding metabolites that form adducts with DNA (Figure 15.8). These agents have been shown to induce liver and bladder cancer in animal models and have been implicated as a cause of bladder cancer in humans. Another aromatic amine compound, 2-acetylaminofluorene (AAF), is metabolized to *N*-hydroxy-AAF and has seen extensive use as a model hepatic carcinogen in rodents to study cancer development and progression.

Inorganic Carcinogens

Several metals compounds have been shown to exhibit carcinogenicity in experimental animals as well as in humans. A summary of some of the more studied of these carcinogenic metals and their organ target for carcinogenicity activity is shown in Table 15.13.

Arsenic In the majority of studies in experimental animals no tumors have been observed or the cancer results were inconclusive for chronic exposure to arsenic. Thus, there is limited or no evidence for the carcinogenicity of inorganic arsenic compounds in experimental animals. In contrast, inorganic arsenic compounds are established human carcinogens. Epidemiological studies of humans exposed to arsenic compounds demonstrated that exposure to inorganic

TABLE 15.13 Carcinogenicity of Metals

	Cancer Results from Animal Studies[a]		Cancer Results in Humans[b]	
Metal	Species	Tumor	Exposure	Tumor Type
Arsenic	Mice, dogs, rats	None found	Cu refinery	Pulmonary carcinoma
			As pesticides	Lymphoma, leukemia
			Chemical plants	Dermal carcinoma
			Drinking water (oral)	Hepatic angiosarcoma
Beryllium	Mice, rats, monkeys	Bone osteosarcoma	None observed	None observed
		Lung carcinoma		
Cadmium	Mice, rats, chickens	Muscle sarcoma	CD refinery	Pulmonary carcinoma
		Testes teratoma		
Chromium	Mice, rats, rabbits	Muscle sarcoma	CR refinery	Pulmonary carcinoma
		Lung carcinoma	Chrome plating	Gastrointestinal carcinoma
			Chromate pigments	
Lead	Mice, rats	Renal carcinoma	None observed	None observed
Nickel	Mice, rats, cats, hamsters, rabbits	Carcinoma	Ni refinery	Pulmonary carcinoma
	Guinea pigs, rats	Lung carcinoma		Nasolaryngeal carcinoma
		Renal carcinoma		Gastric and renal carcinoma
				Possible sarcoma

[a]Studies from animals experimentally treated chronically with the metal.
[b]Results are based on occupational and environmental exposure studies to the metal.

arsenic compounds increases the risk of cancer in the skin, lung, digestive tract, liver, bladder, kidney, and lymphatic and hematopoietic systems. In particular, epidemiological studies have shown a dose–response relationship between arsenic in drinking water and bladder, kidney, lung, and skin cancer. The mechanisms for cancer formation appear to involve the induction of oxidative stress (an increase in the redox of the cell) which can produced DNA damage and/or modify cellular growth and apoptosis. In humans, the latency period for drinking-water-related arsenic carcinogenesis is 30–50 years of exposure.

Beryllium Beryllium and its salts are not mutagenic but bind to nucleoproteins and inhibit enzymes involved in DNA synthesis, resulting in infidelity of DNA synthesis and repair. They also induce gene mutations in cultured cells. An increase in lung tumors in rodents and nonhuman primates exposed to beryllium or beryllium compounds has been reported. Epidemiological studies have similarly shown an increased risk of lung cancer in workers exposed to beryllium or beryllium compounds.

Cadmium Animal studies have shown that cadmium and cadmium compounds induce tumor formation at various sites in several different species of experimental rodents. Tumors were also produced following several different exposure routes and durations. The lesion produced include prostate tumors in rats, testicular tumors in rats and mice, lymphoma in mice, adrenal-gland tumors in hamsters and mice, and lung and liver tumors in mice. Cadmium and cadmium compounds have been classified as known human carcinogens based on evidence of carcinogenicity in humans, including epidemiological and mechanistic information that indicate a causal relationship between exposure to cadmium and cadmium compounds and human cancer. Epidemiological studies of cadmium workers have concluded that exposure to cadmium results in an increase in lung cancer. In addition, although weaker in association, some epidemiological studies have linked cadmium exposure with increased prostate cancer incidence.

Chromium While the metal chromium has several oxidation states, the trivalent (3+) and hexavalent (6+) forms are the most prominent and also the greatest concern regarding human carcinogenicity. Chromium III does not produce cancer in laboratory animals. In contrast, chromium VI has been reported to be carcinogenic in rodents probably through a genotoxic, DNA reactive mechanism. Chronic chromium VI exposure resulted in benign and malignant lung tumors in rats. Similarly in occupational settings, hexavalent Chromium (+6) exposure has been linked to an increase in lung cancer and nasal cancer risk in chromate workers.

Nickel In rats and mice, exposure by inhalation or intratracheal instillation of nickel produced an increase in lung tumors. Injection of soluble nickel salts into rodents produced kidney cancer. Nickel compounds are considered carcinogenic to humans based on the observed elevated risk of death from lung and nasal cancers in nickel workers.

TABLE 15.14 Modes of Action for Selected Nongenotoxic Chemical Carcinogens

Mode of Action	Chemical
Cytotoxicity	Chloroform
	Melamine α2u-Globulin-binding
	d-limonene, 1,4-dichlorobenzene
Receptor-Mediated	
CAR	Phenobarbital
PPARα	Trichloroethylene
	Perchloroethylene diethylhexylphthalate
	fibrates (e.g., clofibrate)
AhR	TCDD
	Polychlorinated biphenyls (PCBs)
	Polybrominates biphenyls (PBBs)
Hormonal	Steroid and peptide hormones
	DES
	Phytoestrogens
	Tamoxifen
Oxidative stress	Acrylonitrile
	1,3-dichlorobenzene

Lead In laboratory studies with rodents, soluble (lead acetate and lead subacetate) and insoluble (lead phosphate, lead chromate) inorganic lead compounds and tetraethyl lead (an organic lead compound) induced kidney tumors as well as tumors of the brain, bone marrow, and lung. Lead exposure has been associated with increased risk to lung, stomach, and bladder cancer in humans.

Nongenotoxic (Epigenetic) Carcinogens

As noted earlier, a number of chemicals that produce tumors in experimental animals act via mechanisms not involving direct binding or damage to genomic DNA. For most nongenotoxic carcinogens, the site of tumor development is many times in tissues where an increased incidence of background, spontaneous tumors is seen. The modes of action for non-DNA reactive carcinogens are diverse and include sustained cytotoxicity, receptor-mediated (e.g., CAR, PPARα, AhR) effects, hormonal perturbation, induction of oxidative stress, and modulation of methylation status (Table 15.14). Each of these potential mechanisms is discussed in the following.

Cytotoxicity Cytotoxicity is a well-established mode of action for a variety of non-DNA reactive chemical carcinogens. Chloroform-induced liver and kidney tumors and melamine-induced bladder tumors are classic examples of chemical carcinogens that function via a cytolethal mode of action. Agents that function through this mechanism produce sustained cell death that is accompanied by persistent compensatory growth. This results in the acquisition of "spontaneous" DNA mutations and allowing mutated cells to accumulate and proliferate. In addition, if the target tissue has already initiated cells, the continual hyperplasia functions at the tumor promotion step of the process. Both of these effects result in tumor formation. It is important to note that a continual necrotic effect with sustained cell divisions is necessary to induce the tumor. Chloroform, for example, has been shown to induce mouse liver tumors only at doses that produce liver necrosis.

The carcinogens *d*-limonene and 1,4-dichlorobenzene induce renal tumors selectively in the male rat, and provide excellent examples of the species, sex, and tissue specificity of non-DNA reactive carcinogens working through a cytotoxicity mechanism. This species and gender specificity is related to the ability of these compounds to bind to α2u-globulin, a protein synthesized selectively in male rat liver. α2u-Globulin is only produced in the male rat and is reabsorbed by the proximal tubules of the kidney. Certain chemicals such as *d*-limonene can to bind α2u-globulin. This binding prevents the breakdown of α2u-globulin, which accumulates in the cells and results in cell death of the tubule cells. Tubule cell death leads to increased cell proliferation of the remaining tubules cells to compensate, which leads to neoplasia.

Receptor-Mediated Nongenotoxic Carcinogens Receptor-mediated nongenotoxic carcinogens work through the activation of receptors in the cell that eventually causes cell proliferation. The constitutive androstane receptor (CAR) is a nuclear receptor that controls the activation of genes involved in the metabolism of xenobiotic substances. In nongenotoxic carcinogenesis, phenobarbital is perhaps the best-known non-DNA reactive compound that functions through the CAR activation mechanism. After exposure to phenobarbital in the liver there is activation of CAR, the induction of P450 enzymes, and an increase in cell proliferation selectively in liver cells that become preneoplastic focal lesions. This results in the formation of neoplasia.

Peroxisome Proliferator Activated Receptor α (PPARα) Another nuclear receptor important in carcinogenesis, the peroxisome proliferator activated receptor α (PPARα) is activated by a number of chemicals such as herbicides, chlorinated solvents (e.g., trichloroethylene and perchloroethylene), plasticizers (e.g., diethylhexylphthalate and other phthalates), lipid-lowering fibrate drugs (e.g., ciprofibrate, clofibrate), and natural products. Most of these agents produce liver enlargement and hepatocellular neoplasia in rodents through non-DNA reactive mechanisms. In contrast, while primates and the guinea pig also have the PPAR receptor, these two species are nonresponsive to the carcinogenic effects of numerous PPAR-activating compounds.

Aryl Hydrocarbon Receptor (AhR) The Aryl Hydrocarbon Receptor (AhR) interacts with compounds such as TCDD

and selective polychlorinated- and brominated-biphenyl (PCBs and PBBs). Activation of the AhR receptor, like the aforementioned two other receptors, has been linked to tumor formation possibly through the tumor promotion stage of the carcinogenic process.

Hormones Finally, hormonally active agents include selective biogenic amines, steroids, and peptide hormones that cause cancer also work through hormone receptors. Thus, a number of non-DNA reactive chemicals induce neoplasms through perturbation of normal hormonal balance. Trophic hormones are known to induce cell proliferation at their target organs. This action may lead to the development of tumors when the mechanisms of hormonal control are disrupted persistently increased hormone levels. Several well-studied examples include the induction of ovarian neoplasms via decreased estradiol and increased LH levels induction of thyroid tumors in rats by phenobarbital-type P450 inducers, and 17β-estradiol exposure leading to an increased incidence of mammary tumors. Evidence that estrogenic agents are carcinogenic in humans comes from epidemiological data on breast and ovarian cancer indicating that those individuals with higher circulating estrogen levels, and those with exposure to the potent estrogenic agent diethylstilbesterol (DES) are at increased risk of cancer development. DES was first shown to induce mammary tumors in male mice following subcutaneous administration of the hormone. DES has been causally associated to the higher incidence of adenocarcinomas of the vagina and cervix in daughters of women treated with the hormone during pregnancy.

Oxidative Stress and Chemical Carcinogenesis Experimental evidence has shown that increases in reactive oxygen in the cell, through either physiological modification or through chemical carcinogen exposure, contribute to the carcinogenesis processes. Reactive oxygen species encompass a series of reactive compounds including the superoxide anion ($\bullet O_2^-$), hydroperoxyl radical ($HO_2\bullet$), hydrogen peroxide (H_2O_2), and the hydroxyl radical ($\bullet OH$), all derived through the reduction of molecular oxygen (Table 15.15). Oxygen radicals can be produced by both endogenous and exogenous sources and are typically counterbalanced by antioxidants. Endogenous sources of reactive oxygen species include oxidative phosphorylation, P450 metabolism, peroxisomes, and inflammatory cell activation (Table 15.16). Reactive oxygen species left unbalanced by antioxidants can result in damage to cellular macromolecules. In DNA, reactive oxygen species can produce single- or double-stranded DNA breaks, purine, pyrimidine, or deoxyribose modifications, and DNA cross-links. Persistent DNA damage can result in either arrest or induction of transcription, induction of signal transduction pathways, replication errors, and genomic instability, all events that are potentially involved in carcinogenesis. The formation of oxidative DNA adducts has been demonstrated to be mutagenic in bacterial and mammalian cells, producing dose-related increases in cellular transformation, and causing G to T transversions in the DNA which are commonly observed in mutated oncogenes and tumor-suppressor genes.

TABLE 15.15 Examples of Reactive Oxygen Species

$O_2 + e^- \rightarrow O_2^{-\bullet}$ (superoxide anion)
$O_2^{-\bullet} + H_2O \rightarrow HO_2^{\bullet}$ (hydroperoxyl radical)
$HO_2^{\bullet} + e^- + H \rightarrow H_2O_2$ (hydrogen peroxide)
$H_2O_2 + e^- \rightarrow OH^- + {}^{\bullet}OH$ (hydroxyl radical)

TABLE 15.16 Reactive Oxygen Species Generation and Removal in the Cell

Cellular Oxidants	
Endogenous generation	Exogenous
Mitochondria	Redox cycling compounds
$O_2^{-\bullet}$, H_2O_2, ${}^{\bullet}OH$	$O_2^{-\bullet}$
Cytochrome P450	Metals (Fenton reaction)
$O_2^{-\bullet}$, H_2O_2	$H_2O_2 + Fe^{2+} \rightarrow OH^- + {}^{\bullet}OH + Fe^{3+}$
Macrophage/inflammatory cells	Radiation
$O_2^{-\bullet}$, ${}^{\bullet}NO$, H_2O_2, OCl^-	${}^{\bullet}OH$
Peroxisomes	
H_2O_2	

Oxidants can be produced via both endogenous and exogenous sources. However, excess production of oxidants results in damage to cellular biomolecules and may impact on neoplastic development.

TABLE 15.17 Carcinogenic Factors Associated with Lifestyle

Chemical(s)	Neoplasm(s)
Alcohol beverage	Esophagus, liver, oropharynx, and larynx
Aflatoxins	Liver
Betel chewing	Mouth
Dietary intake (fat, protein, calories)	Breast, colon, endometrium, gallbladder
Tobacco smoking	Mouth, pharynx, larynx, lung, esophagus, bladder

Causes of Human Cancer

Based on epidemiological and experimental data a number of chemical and processes have been attributed to causing cancer. These include lifestyle-related exposures (Table 15.17), medical therapy (Table 15.18), infection (Table 15.19), as well as occupational exposure (Table 15.20).

Lifestyle Besides drugs and environmental chemicals and lifestyle decisions (Table 15.17) such as smoking and tobacco use, alcohol use, diet, and exposure to sunlight have been linked to cancer development in humans.

TABLE 15.18 Human Carcinogenic Chemicals Associated with Medical Therapy and Diagnosis

Chemical or Drug	Associated Neoplasms
Alkylating agents (cyclophospamide, melphalan)	Bladder, leukemia
Chloramphenicol	Leukemia
Diethylstilbestrol	Vagina (clear cell carcinoma)
Estrogens	Liver cell adenoma, endometrium, skin
Phenacetin	Renal pelvis (carcinoma)
Phenytoin	Lymphoma, neuroblastoma
Thorotrast	Liver (angiosarcoma)

TABLE 15.19 Carcinogenic Viruses, Bacteria, and Parasites

Agent	Neoplasm(s)
Human papilloma viruses (HPVs)	Cervical, penis, anus, vagina, vulva, mouth, throat, head, neck
Epstein-Barr virus (EBV)	Nasopharyngeal cancer, Burkitt's lymphoma
Hepatitis B virus (HBV) and hepatitis C virus (HCV)	Liver
Human immunodeficiency virus (HIV)	Kaposi sarcoma, cervical cancer, non-Hodgkin lymphoma.
Helicobacter pylori	Stomach cancer
Liver flukes	Bile duct cancer
Schistosoma haematobium	Bladder cancer

Tobacco The 1982 U. S. Surgeon General's report stated "Cigarette smoking is the major single cause of cancer mortality [death] in the United States." The Centers for Disease Control and Prevention (CDC) reported that more than 45 million U.S. adults were current smokers in 2010. Smoking accounts for at least 30% of all cancer deaths in the United States and 90% of lung cancer deaths among men and approximately 80% of lung cancer deaths among women are related to smoking. Smoking has been linked to other types of cancer, including cancers of the throat, mouth, nasal cavity, esophagus, stomach, pancreas, kidney, bladder, and cervix, and acute myeloid leukemia. Cigarette smoke is a mixture of chemicals that is produced when tobacco and its additives burn. The smoke is made up of more than 4000 chemicals, including over 60 known to cause cancer in rodents or humans. These include the carcinogens Polycyclic Aromatic Hydrocarbons (PAHs), Aromatic Amines, nitrosamines, Benzene, Formaldehyde, Cadmium, Nickel, Arsenic, and Polonium-210.

Alcohol Use and Cancer Alcohol use has been linked to increased risk of several types of cancer, including cancers of the mouth, throat (pharynx), larynx, esophagus, liver, breast, and colon. The risk of these cancers increases with the amount of alcohol consumed. The combined use of alcohol and tobacco increases the risk of these cancers far more than the effects of either drinking or smoking alone. Alcohol may decrease the ability to repair DNA damage caused by chemicals in tobacco as well as serve as a solvent for some of the harmful materials in the cigarette smoke allowing these agents to get into the cells more efficiently. The amount of alcohol consumed over time, and the form of alcoholic beverage, appear to be important in the cancer risk associated with alcohol use.

TABLE 15.20 Occupational Human Carcinogens

Agent	Industrial Process	Neoplasms
Asbestos	Construction, asbestos mining	Peritoneum, bronchus
Arsenic	Mining and smelting	Skin, bronchus, liver
Alkylating agents (mechloro-ethamine	Chemical manufacturing	Bronchus hydrochloride and bis[chloromethyl]ether)
Benzene	Chemical manufacturing	Bone marrow
Benzidine, beta-naphthylamine	Dye and textile	Urinary bladder
Chromium and chromates	Tanning, pigment making	Nasal sinus, bronchus
Nickel	Nickel refining	Nasal sinus, bronchus
Polynuclear aromatic hydrocarbons	Steel making, roofing, chimney cleaning	Skin, scrotum, bronchus
Vinyl chloride monomer	Chemical manufacturing	Liver
Wood dust	Cabinetmaking	Nasal sinus
Beryllium	Aircraft manufacturing, electronics	Bronchus
Cadmium	Smelting	Bronchus
Ethylene oxide	Production of hospital supplies	Bone marrow
Formaldehyde	Plastic, textile, and chemical	Nasal sinus, bronchus
Polychlorinated biphenyls	Electrical-equipment production and maintenance	Liver

Diet and Physical Activity Each year, about 570,000 Americans die of cancer; it has been estimated that one out of every three cancers in the United States is linked to excess body weight, poor nutrition, or physical inactivity. A healthy weight is important to reduce the risk of cancer. Obese individuals have an increased risk of breast, colon, endometrial, esophageal, and kidney cancer. To reduce cancer risk, most people need to keep their BMIs below 25. A well-balanced diet is also an important part of reducing cancer risk. Vegetables and fruits contain vitamins, minerals, antioxidants, and many other substances that have been shown to reduce cancers of the lung, mouth, esophagus, stomach, and colon. Reducing the consumption of processed meats (hot dogs, bologna, bacon, and deli meats) and red meats like beef, pork, and lamb may help reduce the risk of colon and prostate cancers. While these factors are all related and may all contribute to cancer risk, body weight seems to have the strongest evidence linking it to cancer. Excess body weight contributes to 14–20% of all cancer-related deaths.

Being overweight or obese is clearly linked with an increased risk of many cancers, including cancer of the breast (in women past menopause), colon, endometrium (uterus), esophagus, and kidney. Obesity has also been linked with an increased risk of cancers of the pancreas, gallbladder, thyroid, ovary, and cervix, as well as multiple myeloma, Hodgkin lymphoma, and aggressive prostate cancer.

Skin Cancer Skin cancer is the most common of all cancers. It accounts for nearly half of all cancers in the United States. Three types of skin cancer are seen in humans: melanoma, the most serious type of skin cancer; basal cell cancer; and squamous cell cancer. More than 2 million cases of nonmelanoma skin cancer are found in this country each year. Most skin cancers are classified as nonmelanomas; usually starting in either. These cells are located at the base of the outer layer of the skin or cover the internal and external surfaces of the body. While basal cell or squamous cell cancers are frequently cured if detected early, melanoma is more dangerous due to its rapid invasion and metastasis. Melanoma accounts for over 90% of the skin cancer deaths each year. The causes of skin cancer many a time involve lifestyle choices where there is excessive exposure to ultraviolet (UV) radiation (sunlight).

Viruses, Bacteria, and Parasites Infections from viruses, bacteria, and parasites have been shown to cause or play a role in cancer development. Infections worldwide account for almost 20% of all human cancer. In the United States, less than 10% of all cancers are linked to infectious agents. Several viruses are now known to be linked with human cancer. Human papilloma viruses (HPVs) are a group of related viruses that can cause warts on the skin, mouth, genital organs, and larynx. They are spread by contact (touch), including through sex. HPV infections are very common in people who are sexually active. Certain types of HPV are the main cause of cervical cancer. Most women with cervical cancer shows signs of HPV infection but most women infected with HPV will not develop cancer. HPVs also have a role in causing some cancers of the penis, anus, vagina, and vulva. They are linked to cancers of the mouth, throat, head, and neck too. Epstein-Barr virus (EBV), a herpes virus, is best known for causing mononucleosis and remains in the body throughout life even though most people have no symptoms beyond the first few weeks of infection. Infection with EBV increases a person's risk of getting nasopharyngeal cancer (cancer of the area in the back of the nose) and Burkitt's lymphoma. It may also be linked to Hodgkin lymphoma and some cases of stomach cancer. Hepatitis B virus (HBV) and hepatitis C virus (HCV) cause viral hepatitis, a type of liver infection. In the United States, about 30% of liver cancers are linked to HBV or HCV infection. This number is much higher in some other countries. Of the two viruses, infection with HBV is more likely to cause symptoms, such as a flu-like illness and yellowing of the eyes and skin (jaundice). Human immunodeficiency virus (HIV), the virus that causes acquired immune deficiency syndrome (AIDS), does not appear to cause cancers directly. But HIV infection increases a person's risk of getting several types of cancer. HIV infection has been linked to a higher risk of developing Kaposi sarcoma and invasive cervical cancer. It is also linked to certain kinds of lymphoma, especially non-Hodgkin lymphoma.

Stomach cancer is fairly rare in the United States, but it is the fourth-most common cancer worldwide. Long-term infection of the stomach with *Helicobacter pylori* (*H. pylori*) may cause ulcers. More than half of all cases of stomach cancer are thought to be linked to *H. pylori* infection. Other factors also play a role in whether or not someone develops stomach cancer.

Certain parasitic worms that can live inside the human body can also raise the risk of developing some kinds of cancer. Liver flukes (a flatworm) that have been linked to increased risk of developing cancer of the bile ducts. These infections come from eating raw or undercooked freshwater fish. *Schistosoma haematobium* is a parasite found in the water of the Middle East, Africa, and Asia. Infection with this parasite (an illness called schistosomiasis) has been linked to bladder cancer.

15.6 TEST SYSTEMS FOR CARCINOGENECITY ASSESSMENT

In vivo and *in vitro* experimental systems are used to evaluate if chemicals and pharmaceuticals are carcinogenic (Table 15.21). The tests are defined by duration of exposure, short-term and long-term. The basis for these tests is founded in the mechanism by which carcinogens function. Short-term tests are typically of the duration of days to a

few weeks, while chronic long-term tests usually encompass 6 months to 2 years' exposure to a chemical. These bioassays frequently use bacterial and *in vivo* and *in vitro* mammalian targets.

Short-Term Tests for Mutagenicity

Short-term tests for mutagenicity were developed to identify potentially carcinogenic chemicals based on their ability to induce mutations in DNA either *in vivo* or *in vitro*. A variety of *in vivo* and *in vitro* short-term tests are available to test the potential carcinogenicity of a chemical (Table 15.22). The majority of these tests measure the mutagenicity of chemicals as a surrogate for carcinogenicity. Therefore, while they are usually very predictive of direct- and indirect-acting mutagens (if a metabolic source is provided), these tests routinely fail to detect nongenotoxic carcinogens. Details on the short-term assays for mutagenicity are noted in Chapter 14.

TABLE 15.21 Assays for the Identification of Carcinogenic Potential

Assay	Time Frame
Short-Term	
In vitro mutagenesis assays	Weeks
In vivo DNA reactivity	1–3 months
Subchronic (*in vivo*)	3 months
Long-Term	
Chronic bioassay in animals (rats and mice)	18–24 months in life/1 year data evaluation

Chronic Testing for Carcinogenicity

Two-year studies in laboratory rodents remain the primary method by which chemicals or physical agents are identified as having the potential to be hazardous to humans. The most common rodents used are the rat and the mouse (Table 15.23). Typically, the bioassays are conducted over the life span of the rodents (2 years). Historically, selective rodent strains have been used in the chronic bioassay; however, each strain has both advantages and disadvantages. For example, the National Toxicology Program (NTP) typically uses Fisher 344 (F344) rats and B6C3F1 mice.

In the chronic bioassay, two or three dose levels of a test chemical and a vehicle control are administered to 50 males and 50 females (mice and rats), beginning at 6 weeks of age, continuing throughout their life span. The route of administration can be via oral (gavage), dietary (mixed in feed), or inhalation (via inhalation chambers) exposure. Generally, the maximum tolerated dose (MTD) is used to set the high dose in the chronic study. During the study food consumption and bodyweight gain should be monitored, and the animals observed clinically on a regular basis and at necropsy the tumor number, location, and diagnosis for each animal is thoroughly assessed.

Carcinogenicity Testing in the Skin

The mouse skin model is one of the most extensively studied and used models for understanding multistage carcinogenesis. Originally put forth by Berenblum and colleagues in the 1940s, this model of carcinogenesis is an assay that has been

TABLE 15.22 Short-Term Tests for Mutagenicity

Assay	DNA/Gene Target
Gene mutation assays *in vitro*	
Prokaryote mutagenesis *in vitro* (Ames test, etc.)	Mutations in specific bacterial strains
Mouse lymphoma thymidine kinase (TK)	Mutations in TK
Chinese hamster ovary (CHO) and V79 (HGPRT)	Mutations in HGPRT
Gene mutation assays *in vivo*	
Dominant lethal assay	Death of fertilized egg
Sperm abnormality induction	Microscopically abnormal sperm
Mutation induction in transgenes *in vivo*	
LacZ mouse	Mutations in LacZ gene
LacI mouse	Mutations in LacI gene
LacI rat	Mutations in LacI gene
rpsL mouse	Mutations in rpsL gene
Chromosomal alterations *in vivo*	
Rat bone marrow clastogenesis *in vivo*	Chromosomal aberrations in bone marrow cells *in vivo*
Micronucleus test	Appearance of micronuclei in bone marrow cells *in vivo*
Chromosomal alterations *in vitro*	
Mitotic recombination, or mitotic gene conversion	Conversion of heterozygous alleles to homozygous state
Sister chromatid exchange	Visible exchange of differentially labeled sister chromatids
Primary DNA damage	
DNA repair *in vivo* or *in vitro*	Unscheduled DNA synthesis

TABLE 15.23 Animal Models of Neoplastic Development

	Endpoint
Chronic 2-year bioassay	Tumors in all organs
Tissue specific bioassays	
Liver, mouse	Hepatomas
Lung, mouse	Pulmonary adenomas
Newborn mouse	Neoplasms in liver, lung, lymphoid organs
Multistage models of neoplastic development	
Epidermis, mouse	Papillomas
Liver, rat	Altered hepatic foci
Transgenic mice	
Knockout of p53 tumor suppressor gene ($p53^{def}$)	Tumors in heterozygous animals having normal phenotype
v-Ha-ras	Induced transgene expression in skin leads to papilloma development

TABLE 15.24 IARC Classification of the Evaluation of Carcinogenicity for Human Beings

Group	Evidence
1. Agent is carcinogenic to humans	Human data strong Animal data strong
2A. Agent is probably carcinogenic to humans	Human epidemiology data suggestive Animal data positive
2B. Agent is possibly carcinogenic to humans	Human epidemiology data weak Animal data positive
3. Agent is not classifiable as to carcinogenicity to humans	Human and animal data inadequate
4. Agent is probably not carcinogenic to humans	Human and animal data negative

used to dissect mechanisms of carcinogenesis and is also purported to be a useful intermediate-term cancer bioassay. The number and relative size of papillomas and carcinomas can be quantified as the tumors progress. Both, initiating and promoting activities of chemical carcinogens can be assessed using this model.

Carcinogenicity Testing in the Lung

Strain A mice are genetically susceptible to the development of lung tumors, with spontaneous lung tumors being observed in control animals as early as 3–4 weeks of age with a steady increase to nearly 100% by 24 months of age. The strain A mouse lung tumor assay is sensitive to particular classes of chemicals, such as PAHs, nitrosamines, nitrosoureas, carbamates, aflatoxin, certain metals, and hydrazines.

15.7 CARCINOGEN CLASSIFICATION

Based on the results of the bioassays noted earlier, regulatory agencies will assess the potential carcinogenic risk to humans. Multiple regulatory agencies exist based on national, regional, and target audience (pharmaceuticals, occupational, environmental) that produce risk assessment profiles. For determining risk, the evaluation usually encompasses epidemiological information, experimental animal data, and *in vitro* data from the required genotoxicity and carcinogenicity bioassays. Two widely used classifications include those used by the International Agency for Research on Cancer (IARC) (see Table 15.24) and the USEPA (Table 15.25).

The IARC approach assigns the chemical or mixture to one of five groupings based upon strength of evidence for the agent's possible, probable, or definite carcinogenicity to humans. In the Group 1 classification, the agent or mixture is classified as definitely carcinogenic to humans. This designation is based on strong human epidemiological evidence supported by positive cancer induction in experimental animals. This classification leaves no doubt that the agent is a hazard for cancer risk to humans. The second grouping is designated as Group 2. Classification into 2A occurs when the agent is probably carcinogenic to humans based on a strong tumor response in test animals when exposed, and suggestive (but not definitive) epidemiological data in humans. Similarly, classification into a Group 2B designation occurs when there is strong animal data for carcinogenic activity and weak or absent epidemiological data to support carcinogenic risk. In Group 3, the data for both animal and human epidemiological studies are nebulous or weak and thus the carcinogenic risk is not classifiable. And in the last group, Group 4, the data obtained for both animal and human epidemiological studies are strong and show no cancer-inducing activity.

TABLE 15.25 USEPA Cancer Guidelines Descriptors

Carcinogenic to humans
Strong evidence of human carcinogenicity, including convincing epidemiologic evidence of a causal association between human exposure and cancer
The mode(s) of carcinogenic action and associated key precursor events have been identified in animals, and there is strong evidence that the key precursor events in animals are anticipated to occur in humans
Likely to be carcinogenic to humans
Weight of the evidence is adequate to demonstrate carcinogenic potential to an agent in animal experiments in more than one species, sex, strain, site, or exposure route, with or without evidence of carcinogenicity in humans
Suggestive evidence of carcinogenic potential
The weight of evidence is suggestive of carcinogenicity; a concern for potential carcinogenic effects in humans is raised, but the data are judged not sufficient for a stronger conclusion
Inadequate information to assess carcinogenic potential
Available data are judged inadequate for applying one of the other descriptors
Not likely to be carcinogenic to humans
This descriptor is appropriate when the available data are considered robust—there is no basis for human hazard concern, evidence in both humans and animals that the agent is not carcinogenic

TABLE 15.26 Mode-of-Action Definitions and Framework

Mode of Action
Those key events and processes, starting with the interaction of an agent with a cell, through functional and anatomical changes, resulting in cancer or other health endpoints
Key Event
Observable precursor step that is itself a necessary element of the mode of action or is a biologically based marker for such an element
Mode of Action, Criteria
Summary description of the hypothesized mode of action
Identification of key events
Strength, consistency, specificity of association
Dose–response concordance
Temporal relationship
Biological plausibility and coherence
Consideration of the possibility of other MOAs
Questions
Is the mode of action sufficiently supported in the test animals?
Is the mode of action relevant to humans?
Which populations or life stages can be particularly susceptible to the mode of action?

The IARC approach utilizes a transparent examination of the data by experts and produces a series of monographs that describe the methodology for the evaluation of specific chemicals and the rationale for their final classification.

Similar classifications exist for the U.S. EPA, the U.S. Food and Drug Administration, and the European Community (EC). The classification of agents with regard to human carcinogenicity can many a time be very difficult in particular, when animal data and/or epidemiological data in humans are inconclusive, inconsistent, or confounded.

The new U.S. EPA Guideline for Cancer Risk Assessment (2005) uses descriptors for defining the relative carcinogenic risk to humans (Table 15.25). These descriptors include: carcinogenic to humans; likely to be carcinogenic to humans; suggestive evidence of carcinogenic potential; inadequate information to access carcinogenic potential; and not likely to be carcinogenic in humans. The EPA Guidelines take into account the understanding of the mode of carcinogenic action and associated key precursor events needed for the cancer to form (Table 15.26). Central to the U.S. EPA Guidelines for Cancer Risk Assessment is the utilization of the mode-of-action framework to define the key events in rodents and assessment of whether those same key events and mode-of-action can occur in humans. This approach is similar to that developed by the International Program on Chemical Safety and by panels in the International Life Sciences Institute.

15.8 SUMMARY

The induction of cancer by chemicals has been scientifically studied for over three centuries. Initially using epidemiological observations followed by laboratory-based experimental studies, it is now well-established that certain chemicals and

physical agents (radiation) found from human lifestyle choices, occupational exposure, and environmental exposure can produce or contribute to cancer development in humans. Carcinogenesis is a multistage event that requires changes to the genomic DNA (mutation) as well as epigenetic modifications producing uncontrolled growth of the mutated cell. While the chemical carcinogenesis process requires multiple steps to occur, three definable and demonstrable steps, initiation, promotion, and progression, have been established. Chemical and physical carcinogens can produce their effects on one or more of these steps. Modification and mutation of specific groups of normal cellular genes, Proto-Oncogenes and Tumor-Suppressor Genes, are important in the tumor formation and can be targeted by chemicals and radiation. The development and progression of cancer, like all chronic diseases, involves gene–environment interaction. A role for nuclear polymorphisms in genes involved in carcinogen metabolism and cellular DNA repair/detoxification has recently been linked to human cancer risk and susceptibility. The mechanism for the induction of cancer by chemicals appears to involve either genotoxic (DNA reactive) or non-genotoxic (epigenetic) mechanisms. Most DNA reactive compounds require metabolism by P450 cytochromes to produce the ultimate mutagenic form. In contrast, a large grouping of chemicals that induce tumors in experimental models appear to work through receptor, cytotoxic or oxidative stress mechanisms, and function in increasing the target cell growth. The identification of cancer-causing agents is of great importance and has led to the development of specific defined assays using *in vitro* and *in vivo* bioassays to help with detection. Results from these bioassays are then evaluated and human risk assessment to the chemical in question is made.

SUGGESTED READING

Anand P, Kunnumakkara AB, Kunnumakara AB, Sundaram C, Harikumar KB, Tharakan ST, Lai OS, Sung B, Aggarwal BB. Cancer is a preventable disease that requires major lifestyle changes. Pharm Res 2008;25(9):2097–2116.

Butterworth B, Slaga TJ. *Nongenotoxic mechanisms in carcinogenesis. Banbury Report*. Cold Spring Harbor: Cold Spring Harbor Press; 1987.

Cerutti PA, Trump BF. Inflammation and oxidative stress in carcinogenesis. Cancer Cells 1991;3(1):1–7.

Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011;144(5):646–674.

Hsu C. *Study Guide for: Cancer Risk Assessment: Chemical Carcinogenesis, Hazard Evaluation, and Risk Quantification.* Content Technologies; 2013.

International Agency for Research on Cancer (IARC). *Monographs on the Evaluation of Carcinogenic Risks to Humans. Volume 100C. A Review of Human Carcinogens: Arsenic, Metals, Fibres, and Dusts*; 2012.

Klaassen C. *Casarett & Doull's Toxicology: The Basis Science of Poisons*. 8th ed. New York: McGraw-Hill Professional; 2013.

Klaunig JE, Ruch RJ. Role of inhibition of intercellular communication in carcinogenesis. Lab Invest 1990;62(2): 135–146.

Klaunig JE, Wang Z, Pu X, Zhou S. Oxidative stress and oxidative damage in chemical carcinogenesis. Toxicol Appl Pharmacol 2011;254:86–99.

Kleinsmith LJ. *Principles of Cancer Biology*. San Francisco: Pearson Benjamin Cummings; 2006.

National Research Council. *Health Risks from Exposure to Low Levels of Ionizing Radiation: BEIR VII—Phase 2*. Washington: National Academies Press; 2006.

Pitot HC. The molecular biology of carcinogenesis. Cancer 1993;72:962–970.

Tannock I, Hill R, Bristow R, Harrington L. *The Basic Science of Oncology*. 4th ed. New York: McGraw-Hill; 2005.

Wakeford R. The cancer epidemiology of radiation. Oncogene 2004; 23(38):6404–6428.

Warshawsky D, Landolph JR Jr. *Molecular Carcinogenesis and the Molecular Biology of Human Cancer*. Boca Raton: Taylor and Francis Group; 2006.

Weisburger JH, Williams GM. Bioassay of carcinogens: in vitro and in vivo tests. In: Searle CE, editor. *Chemical Carcinogens*. Vol 2. Washington: American Chemical Society; 1984. p 1323–1373.

16

PROPERTIES AND EFFECTS OF METALS

David B. Mayfield, Ari S. Lewis, Lisa A. Bailey, and Barbara D. Beck

Investigations into the effects of metals on the human body have been of interest to physicians, scientists, and in some cases, those of a criminal bent, for millennia. The ancient Romans, Egyptians, Chinese, and other civilizations were aware of the effects of multiple metals, such as lead, antimony and arsenic. Metals were understood to elicit various toxicological reactions, particularly from high-level acute exposures eliciting readily apparent symptoms. Metals were also perceived to have various therapeutic abilities, such as curing coughs and asthma (arsenic), treating hangovers and other overindulgences (antimony), and curing tuberculolis (lead). The more sinister side of metals is another important part of their history. This is especially relevant with arsenic, a metal (actually a metalloid) with a particularly notorious history. Arsenic was desirable as a homicidal tool because the symptoms of acute arsenic poisoning are relatively general, making determination of cause of death difficult. However, the development of the Marsh test in the nineteenth century for detection of arsenic in body fluids reduced, but did not eliminate, its use for this purpose. As recently as 2003, arsenic was used intentionally in a poisoning incident in the United States in the state of Maine.

The toxicity of metals remains a robust area for scientific inquiry today. The awareness of exposure sources has expanded over time and includes pharmaceutical products, environmental media, consumer products, and workplace. Interest in types of metals has also expanded to include a broad range, some of which, such as the rare earth elements (REEs), that is, the lanthanides, have only been studied recently. New effects of certain metals already relatively well studied are being identified, for example, gastrointestinal (GI) tumors in rodents ingesting hexavalent chromium in drinking water. And methodologies for investigating metals have advanced well beyond descriptive studies of clinical effects, reflecting greater understanding of toxicological mechanisms at the cellular and molecular levels. Consider, for example, the use of biomarkers to evaluate exposures to and effects of metals.

In this chapter, we

- Provide an overview of general considerations for understanding the fate, including metabolic transformations and changes in valence state, and transport of metals in the environment, including air, water, and soil
- Describe basic toxicological considerations for metals, including absorption, distribution, metabolism, and excretion. The aim of these general discussions is not to generalize across types of metals (which is, of course, scientifically inappropriate) but to identify important features that must be considered when evaluating exposures to and toxicity of different metals and of different forms of the same metal
- Provide brief overviews of 13 metals from A (arsenic) to Z (zinc), along with discussion of the lanthanides

16.1 GENERAL CHARACTERISTICS OF METALS

Metals are naturally occurring elements that are present throughout the environment. Many metals are essential nutrients and are required at some minimum level for good health; therefore, the distinction must be made between necessary minimal exposure and toxic overexposure. Organisms, including humans, have evolved detoxifying mechanisms to cope with potentially harmful levels of both essential and nonessential metals since exposure to metals

Principles of Toxicology: Environmental and Industrial Applications, Third Edition. Edited by Stephen M. Roberts, Robert C. James and Phillip L. Williams.

occurs regularly in the environment. It is when the frequency, magnitude, or duration of exposure exceeds the detoxification capacity that metal toxicity may become a concern.

Metals and metalloids are typically found in a solid state at room temperature, except for elemental mercury, which exists in the liquid phase. Metals can transform from solid to liquid or liquid to gas depending on their respective melting and boiling points. Due to their physical state, exposure to metals via inhalation primarily occurs in particulate form or as a dust or fume. Exceptions are elements that occur in gaseous form at certain temperatures (e.g., mercury vapor, arsine gas (AsH_3), and stibine (SbH_3)). Metals are also common in foods, soils, and tissues complexed with organic matter or in dissolved form complexed with ionic compounds (e.g., chlorides or sulfates). These metal complexes account for metal ingestion exposures via water, soil, and diet.

Metals, like all elements, cannot be broken down or degraded and are therefore persistent in the environment. They can, however, form compounds with other elements and chemicals. The primary feature that distinguishes metals from nonmetals is their ability to readily lose electrons in solution and form positively charged ions (cations). The cations can form strong ionic bonds (complexes) with nonmetals. While some important metals may be found in their elemental state (e.g., mercury) in the environment, most metals are found in a wide variety of complexes with other substances. The interactions of metal cations with their surroundings drive both their mobility in the environment and their biological effects. The various metal complexes, even with the same metal cation, may differ dramatically in their chemical and toxicological properties (e.g., elemental mercury vs. methyl mercury).

While humans possess effective means for detoxifying and excreting metals at exposure levels normally encountered in the environment, some human activities lead to increased exposure to metals. In occupational situations, protection against exposure to metals is controlled through the use of improved work practices and protective equipment or by simply reducing exposure. Exposure to metals in the general population has also increased. Mining, processing, commerce, and waste disposal have all increased the movement of metals in the environment, both directly and indirectly, and contribute to higher background levels of exposure in general. In this chapter, properties and potential effects of both occupational and environmental exposure to metals are discussed.

16.2 ENVIRONMENTAL AND OCCUPATIONAL EXPOSURE

Natural State of Metals in the Environment

Metals are naturally present in the earth's crust, and thus human exposure from natural sources is inevitable. Metals are found in a range of concentrations in soils, sediments, surface and ground waters, and air. Table 16.1 lists the typical levels of some metals found in the environment. It is important to remember that these levels may vary considerably due to geologic characteristics, land use (industrial vs. rural vs. urban), and the effects of human activities (e.g., mining operations and hazardous waste facilities). Thus, the levels reported in Table 16.1 are only ranges for comparative purposes and are not expected to represent background levels in all areas.

Metals move through the environment as part of natural biogeochemical pathways: dissolved from rocks; transported in soils, water, and air; taken up from soil and water by plants; passed up the food chain; and released back to the environment through excretion and decay of dead organisms and often going through various transformations in the process. Metals may form a variety of compounds by combining with available negatively charged groups, such as chlorides, sulfates, nitrates, carbonates, and acetates. The metal complexes can differ dramatically from the uncomplexed metal in their solubility and their absorption by and distribution within living organisms. Therefore, the toxicological properties of metals and their complexes can vary significantly.

Transformations Speciation is the primary characteristic that determines a metal's potential for exposure and its resulting toxicity. A metal may exist in its elemental, uncharged, form and this state is often designated by a "0" representing "zero charge" (i.e., Hg(0) or Hg^0 for elemental mercury). This is typically the form of objects that we think of as being metallic. However, a property common to all metals is the tendency to ionize in solution, giving up one or more electrons to become a cation or a compound with a net positive charge (e.g., for cadmium, $Cd \rightarrow Cd^{2+} + 2e^-$). Metals may exist in a variety of ionization states, or valences, that differ in the number of missing electrons and therefore exhibit various net positive charges. The standard designation for the first ionization state of many metals is the name of the metal with an *ous* suffix, such as *mercurous* ion for Hg^+ and *cuprous* ion for Cu^+. The next level of ionization is sometimes designated with an *ic* suffix, such as *mercuric* for Hg^{2+} and *cupric* for Cu^{2+}. Ionization is not a uniform process across all metal types, and a metal may give up two electrons in its first ionization state (e.g., for tin, the *stannous* ion is Sn^{2+}) or even three electrons (e.g., trivalent chromium is Cr^{3+} or arsenite is As^{3+}). Other ionization states may exist for some metals, but usually the number of states most commonly encountered under normal conditions is limited to two or three. The degree to which this process is carried out will, in turn, affect various metal characteristics, including its toxicity.

Metal speciation is heavily dependent on physical parameters such as pH (the concentration of H^+ ions), oxidizing or reducing conditions (often measured as reduction potential (Eh) or oxidizing–reducing potential), and temperature. The

TABLE 16.1 Typical Environmental Concentrations of Metals

	Soil		Domestic Groundwater		Food		Air	
	10th	90th	10th	90th	Typical Range		Typical Range	
	(mg/kg)		(μg/l)		(mg/day)		(ng/m^3)	
As	1.8	10	0.11	7.53	0.05		20	30
Be	0.5	2.1	<1	<1	No information		0.03	0.2
Cd	<0.1	0.5	<1	<1	0.03		1	40
Cr	10	52	0.17	4	0.06		10	30
Cu	5	23	0.17	12.3	0.9	1.3	5	200
Pb	9.6	27	0.01	1.09	Highly variable		Highly variable	
Mn	163	1140	0.05	172.39	1	10	20	
Mo	0.26	1.48	0.13	6	No information		No information	
Hg	<0.02	0.08	nm	nm	0.0035		6	20
Ni	5.6	31	0.1	3	0.17		7	12
Se	<0.2	0.7	0.13	3.02	0.071	0.152	No information	
Tl	0.2	0.7	<1	<1	No information		No information	
Zn	22	90	1	99.86	7	16.3	100	1700

Source: USGS (2005, 2009); ATSDR (2010).
Soil and groundwater concentrations for the 10th and 90th percentile concentrations are given.
Food and air concentrations are given as typical ranges, as defined by the ATSDR, and singular estimates of average intakes for adults.
For lead, the air and food sources are highly dependent on local conditions, so the ATSDR did not present a typical range.
nm, nanometer.

presence of other charged particles that can bind with the metals (known as ligands) in the air, soil, or water is another strong determinant if the metal will remain as a free cation or form a complex. These parameters determine how and at what rates metals may move through the environment, which must be considered when assessing the primary routes of exposure and potential concentrations during exposure. Metal cations are typically more mobile than their elemental forms, but it is difficult to generalize how other external factors influence metal mobility since a single factor might increase the mobility of one metal while dramatically reducing that of another. For example, a common ligand in groundwater, sulfate, enhances the solubility (and therefore mobility) of cadmium and zinc, while barium sulfate is nearly insoluble and is not readily transported.

An example of the degree to which these chemical and physical factors affect human exposure to metals can be illustrated by copper. Copper has an especially high affinity for organic ligands. In aquatic systems, copper is ionized and readily binds to organic matter, settles out of the water column, and is unavailable to most organisms. Likewise, copper cations bind to organic matter in soil and tend to accumulate in the upper layers of soil when applied at the surface, as would occur from deposition from air or as runoff from roadways. This surface accumulation decreases the likelihood of oral exposure to copper through contaminated groundwater but increases the likelihood of exposure through soil contact.

Chemical and physical attributes (e.g., chemical species, exposure matrix, and particle size) of metals will also affect absorption into the bloodstream. That is, a metal's bioavailability depends on both the innate properties of that metal as well as how it interacts with environmental media. For example, arsenic is almost 100% bioavailable in water, but in a soil matrix, absorption into the body will be significantly less. The extent of bioavailability is often limited by binding to ligands in the soil and will be influenced by a variety of soil characteristics including the iron and organic matter content as well as the pH. These factors have the potential to reduce the oral bioavailability of arsenic to 50% or less. In the case of metal exposure through inhalation, smaller particles typically result in greater bioavailablity, in part because smaller particles will penetrate deeper into the respiratory tract. In addition to environmental factors, the bioavailability of a metal can also be influenced by host factors, such as gender, age, nutritional status, and other health conditions.

Speciation of metals and conversion from one form to another in the environment may also occur as a result of biological action, termed *biotransformation*. The metabolic processes of many microorganisms can transform metal compounds to more or less toxicologically important forms. In particular, the methylation of metals often results in complexes that are more easily absorbed and therefore potentially more toxic. For example, in contaminated aquatic systems, the transformation of ionic mercury to methyl mercury and to dimethyl mercury is driven by anaerobic microorganisms. Likewise, inorganic lead may be biotransformed by some bacteria to tetraethyl lead, an organic compound that is more easily transported across biological membranes. Biological processes can also work to make metals less available, such as when metals are absorbed

from the soil and groundwater by plants and become incorporated into their tissues. For example, the methylation of inorganic arsenic reduces toxicity because methylated forms of arsenic are poorly absorbed into human cells.

Anthropogenic Exposures In addition to natural biogeochemical cycles, humans play a large role in the mobilization, transformation, and transportation of metals in the environment. Industrial activities, such as mining and manufacturing, often concentrate metals from natural minerals and this increases the potential for human exposure. Some of the concentrated metals may be incorporated into new products and become widely distributed, while others are simply discarded as waste. Human activities can also indirectly alter natural biogeochemical cycles and change how metals move through the environment. A potential by-product of mining is the release of high levels of acid, which mobilizes naturally present metals, into local streams and is known as acid mine drainage. Occupational exposures typically occur during the initial processing of minerals or in the manufacturing of new products (e.g., zinc used for galvanizing steel and lead released during the smelting process). Exposure assessment for the general population must consider the life cycle of a product from its origin in industry, the patterns of use, and ultimately the disposal of both manufacturing and product wastes.

Exposures to high levels of metals are usually of greatest concern for workers in industries where metals are commonly used. These include mining, processing and smelting, manufacturing, and waste disposal operations. The metals of concern in occupational settings are dependent on the specific industry. For instance, sheet metal workers may be exposed to copper and aluminum dust particles, while lead ore smelting workers may be exposed to releases of lead, arsenic, and cadmium. Although workers are often exposed to higher metal concentrations than the general public, health and safety regulations are designed to maintain exposure at safe levels through improved work practices and personal protective equipment, when necessary. Because workplace exposure is often confined to a specific site with a specific population at risk, routine monitoring of exposure can be performed and corrective action can be taken when necessary to maintain safe limits.

People living near industrial areas, waste sites, mining operations, or smelters may be exposed to higher-than-background levels of metals in air, drinking water, and soil. Industrial contamination of local environments has resulted from accidental releases as well as improper management practices. One of the first documented cases of widespread illness from incorrect waste disposal occurred in Japan and underscored the dangers of mercury. Beginning in the 1930s, industrial wastes containing methyl mercury were discharged into Minamata Bay, where it bioaccumulated in fish and shellfish. By the late 1950s, severe neurological impairment, sometimes leading to death, was observed in both adults and children and was later linked to the consumption of mercury in seafood. The neurological symptoms described there later became known as *Minamata disease*.

Increased exposure to metals can also occur through the normal use of products. The use of lead in gasoline and paints resulted in perhaps the most well-known example of increased metal exposure to the general population. While lead was added to both products to improve special characteristics, this had the unintended consequence of increasing the amount of lead in dust and soils to levels far above background concentrations in urban areas. The exposures to airborne lead particles from leaded gasoline as well as dust and soil at homes with lead paint contributed to an overall increase on the lead body burden in the general population and resulted in significantly higher blood lead levels (BLLs) in young children. The discontinuance of lead as an additive has dramatically decreased the impact of these exposure routes, and BLLs have decreased throughout the population.

Exposures from Naturally Occurring Metals In some areas, naturally occurring metals found in soils, sediments, surface and ground waters, and air may be a significant source of exposure. For example, naturally occurring sources of uranium have caused high levels of radioactivity in groundwater in the Floridan aquifer system in southern Georgia. The level of radioactivity, above acceptable limits for drinking water, has resulted in the need for alternative drinking water sources. Some areas of the United States and other parts of the world have groundwater with elevated arsenic of geologic origin. China, Bangladesh, and Taiwan are among the areas of the world with the highest levels of naturally occurring arsenic in groundwater. Arsenic-containing water in areas of these countries can pose a serious public health concern, increasing the risk of many types of cancer and noncancer effects (such as skin lesions, cardiovascular disease, and respiratory illness). Although not a public health concern, another naturally occurring metal with significant human exposures is aluminum. Aluminum enters the environment through the weathering of rocks and minerals, and, because of its abundance in the earth's crust, natural weathering processes far exceed releases into air, water, and land from anthropogenic sources. Aluminum is also found naturally in many foods.

16.3 GENERAL PROPERTIES OF ABSORPTION, DISTRIBUTION, METABOLISM, AND EXCRETION OF METALS

Absorption and Distribution

For a metal to exert a toxic effect on a human or other organism, there must be exposure and it must gain entry into the body. There are three main routes of absorption:

inhalation, oral, and dermal. The significance of the route is metal-specific. For instance, copper is highly soluble in water so that ingestion (oral exposure) is a common pathway of exposure. Ingestion is also an important exposure pathway of mercury, often through the consumption of fish or other animals that bioaccumulate mercury to a high level in their tissues. However, the elemental form of mercury has a high vapor pressure so that inhalation exposure also may be significant in certain situations. Dermal exposure to metals may cause local effects, but it is rarely a significant consideration from an absorption perspective. Absorption through dermal barriers occurs primarily through passive diffusion, which is influenced by molecular weight, lipid solubility, temperature, pH of the substance, concentration, and integrity of the skin. Contact dermatitis and allergic skin reactions may occur in some individuals from dermal exposure to certain metals (e.g., beryllium, nickel, chromium, and cobalt).

The route of absorption may also influence the subsequent distribution of the metal within the body and, thus, may affect its metabolism, potential toxic effects, and excretion. For example, while elemental mercury is effectively absorbed via inhalation into the bloodstream where it can cross the blood–brain barrier and cause neurological effects, elemental mercury is not well absorbed through the oral or dermal route and is essentially nontoxic via these pathways. In contrast, organic forms of mercury are well absorbed through all routes and have even resulted in death via dermal exposure.

Metabolism and Storage

Once metals have entered the bloodstream, they are available for distribution throughout the body. The rate of distribution to organ tissues is determined by blood flow to the particular organ. The eventual distribution of a metal compound is largely dependent on the ability of the compound to pass through cell membranes, coupled with its affinity for binding sites. Metals are often concentrated in a specific tissue or organ (e.g., lead in bone and cadmium in liver), and their metabolism within the body usually involves binding to proteins, such as enzymes, or changes in their speciation. Metals also may bind to other substrates and alter the bioavailability of important cell constituents.

Excretion

The simplest method of excreting metals is to exclude them before they achieve significant absorption. For instance, inhaled metal vapors may be immediately expired during subsequent breaths, while inhaled particles may be trapped in the mucous and expelled by the cough reflex. Vomiting is a common reflex in response to orally ingested toxicants, including metals. Metals may also be excreted in the sweat and saliva or incorporated into growing hair and fingernails (through complexes with sulfhydryl groups). The body's attempt to excrete excess mercury and lead in the saliva often results in visible "lead lines" along the gums. Urine and fecal excretion, however, remains the primary route of excretion for most ingested metals and are also important following inhalation and dermal exposure.

Usually, excretion consists of a combination of these pathways, which may differ according to the route of exposure and the speciation of the metal. Inhaled elemental mercury vapor, for instance, is excreted in the urine, feces, and breath, while ingested elemental mercury is primarily excreted in the feces. Methyl mercury is excreted very slowly, primarily in the feces after transformation to inorganic forms of mercury.

Additionally, a common means by which the body defends itself against metal poisoning is by producing nonenzymatic proteins, which bind to and inactivate the metal, then transport it out of the body via the excretory system. Lipoproteins located within renal lysosomes appear to function in this way, thus serving as some protection for the kidneys, which are particularly vulnerable to metal damage. Metallothioneins (a class of sulfur-containing, metal-binding proteins that facilitate metal excretion) display a substantially increased expression in response to cadmium, mercury, zinc, and several other metals.

Chelation

A number of antidotes to metal poisoning have been developed on the basis of knowledge of natural detoxification mechanisms. Chelation therapy involves systematic treatment of the patient with a chelator, defined as a molecule with several electronegative groups able to form coordinate covalent bonds with metal cations and thus render the metal unavailable and inactive. An example of a chelator is ethylenediaminetetraacetic acid (EDTA), a flexible molecule with four binding sites capable of nonspecifically binding metal ions and escorting them to the kidney for excretion. Calcium EDTA is particularly effective against lead poisoning. British anti-Lewisite (dimercaprol; BAL) is another chelator that was developed during World War II as an antidote against a form of arsenic gas (British Lewisite) and has since found application in therapy for chromium, nickel, cobalt, lead, and inorganic mercury poisoning. Dimercaptosuccinic acid (DMSA) is a water-soluble chelating agent similar to BAL, which is used to treat heavy-metal poisoning. DMSA has fewer side effects than BAL.

Although there are many benefits to chelation therapy, it must be used with discretion since chelators have been shown to result in redistribution of toxic metals, hepatotoxicity and nephrotoxicity, headache, nausea, and increased blood pressure. Chelation therapies are nonspecific in nature and, in addition to the target metal (e.g., arsenic and lead), can bind to other essential trace metals (e.g., calcium, copper,

and zinc), leading to deficiency of these essential metals and unintended health effects if their use is not monitored appropriately. Some chelators have been shown to increase the biologic activity and toxicity of the metal. For example, EDTA is not able to shield the surface of the iron ion (Fe^{3+}) but instead forms an open basket–like complex, resulting in increased catalytic capacity for Fe^{3+} to generate oxidative stress. BAL also has been shown to worsen the effects of selenium and methyl mercury intoxication.

Physiologically Based Pharmacokinetic Models

Physiologically based pharmacokinetic (PBPK) models use empirically based mathematical descriptions of uptake and distribution of toxic substances to quantitatively describe relationships among biological processes. They are used to estimate internal dose of a putative toxic moiety in a target tissue following various combinations of exposure route, dose level, and test species. PBPK models refine our understanding of complex dose relationships, such as external exposure concentration for a given species versus tissue dose of the toxicant and tissue dose versus observed adverse effect of the toxicant. In addition, PBPK models can be used to extrapolate pharmacokinetics from high to low doses, from one exposure route to another, from one species to another, and from one subpopulation to another within a species.

PBPK models have been developed for a number of metals (e.g., manganese, arsenic, chromium, lead, mercury, and nickel) and are often used in risk assessment of these metals to predict tissue dose based on a given exposure and to extrapolate safe environmental exposure levels for humans. For example, PBPK models have recently been developed for manganese, first for rodents and nonhuman primates and more recently for humans, to determine target tissue dosimetry in various regions of the brain that are associated with neurotoxicity following inhalation exposure to high concentrations of manganese.

16.4 BIOMARKERS

Metal exposure and effect can be measured through the use of biomarkers. Biomarkers can be broadly categorized (as defined by the International Program of Chemical Safety) as follows:

- Biomarker of exposure: an exogenous substance or its metabolite or the product of an interaction between a xenobiotic agent and some target molecule or cell that is measured in a compartment within an organism.
- Biomarker of effect: a measurable biochemical, physiological, behavioral, or other alteration within an organism that, depending on the magnitude, can be recognized as being associated with an established or possible health impairment or disease.
- Biomarker of susceptibility: an indicator of an inherent or acquired ability of an organism to respond to the challenge of exposure to a specific xenobiotic substance.

Examination of exposure to metals has typically been evaluated through analysis of urine, blood, hair, or other biological tissues. For example, blood lead monitoring was key to identifying and tracking exposure to lead in gasoline. As the use of lead decreased in the 1980s, a concomitant reduction in BLLs was observed in the U.S. population. Measurement of mercury concentrations in hair has routinely been used to document exposure to methyl mercury and generally corresponds to roughly 250 times the concentration measured in blood. Arsenic tends to accumulate in keratin-rich tissues, and exposure may manifest as discolored bands or *Mees' lines* in fingernails and toenails. Finally, elevated exposure to silver compounds may result in a gray or blue–gray discoloration of the skin, termed argyria.

Biomarkers of effect can be measured through biochemical, molecular, or behavioral indices. For example, impaired kidney function resulting from cadmium exposure may be identified by increased levels of serum proteins in the urine. Biomarkers for cancer development include tests for chromosomal alterations or changes in DNA, RNA, and protein expression. BLLs, in addition to being used as a determinant of exposure, are also used to monitor for cognitive impairment as BLLs at or below 10 μg/dl have been linked to deficits in IQ. Interpretation of effect biomarkers can be difficult because many biomarkers may not be specific to a single metal or chemical.

Recent advances in toxicogenomics and molecular epidemiology are providing tools to identify genetic markers that signify susceptibility to a toxicant. For example, the protein metallothionein is important for the detoxification of metals. When normal levels of metallothionein are diminished due to metal exposure, adverse health effects may occur. Thus, measuring the gene expression of proteins such as metallothionein can provide a biomarker of susceptibility. Examination of genetic polymorphisms and other genetic biomarkers will enhance monitoring and diagnosis of environmentally related diseases.

16.5 GENERAL MECHANISMS OF TOXICITY

The interaction of metals with organs, tissues, cells, and cellular components can result in toxic effects. This may occur by a number of mechanisms that are specific to each metal. However, common toxic mechanisms of action for metals include the following categories:

- Direct effects: Direct toxic effects may occur when the primary compound or metal directly damages or disrupts cellular function or cellular structure without biotransformation. For example, hexavalent chromium is

corrosive and may cause sores on the skin or ulceration of the nasal septum.

- Indirect effects: Metal anions and cations may bind with cellular components, enzymes, or other substrates altering biological function. For example, arsenic and mercury have an affinity for thiol groups (sulfur-containing molecules) and therefore may bind to and inhibit enzymes or alter the structure or function of proteins. Lead exposure can disrupt voltage-/Ca^{2+}-activated K^+ channels resulting in neuronal and muscular inhibition.
- Substitution: Some metals are chemically similar and may bind to proteins or enzymes or accumulate in tissues in a similar manner as essential elements. For example, lead and REEs share similar properties as calcium and are known to accumulate in bones and alter the cellular binding of calcium.
- Metal imbalance: Overexposure to some metals may disrupt normal biological homeostasis and result in the retention or depletion of other essential metals. In some individuals with Wilson's disease (a genetic defect altering copper regulation), copper can accumulate in tissues resulting in toxic effects. In these same individuals, treatment with molybdenum can be used to reduce copper levels.

16.6 ESSENTIALITY VERSUS TOXIC EFFECTS

Essential nutrients are vitamins and trace elements that are critical for normal biological function (e.g., calcium, copper, sodium, and zinc). While nonessential metals are those elements that have no recognized beneficial biological role (e.g., cadmium, mercury, and lead). Some metals are essential for some animals and plants (e.g., boron, nickel, and silicon). Essential metals are components of proteins and enzymes that participate in homeostatic regulation of cellular function and metabolism, hormone levels, and growth and development (see Table 16.2). Essential metals have a "U"-shaped dose–response curve where low levels (i.e., nutrient deficiency) and high levels of metal exposure can disrupt the regulatory capacity of the organism. Recommended dietary reference intakes have been established for many essential metals (Table 16.2), and these represent a daily dietary intake sufficient to meet the nutrient requirements of nearly all healthy individuals. Tolerable upper intake levels (TULs) and chronic oral reference doses (RfDs) represent daily intakes that if exceeded can present a health risk from exposure to the metal (Table 16.2).

Essential metals can be toxic when intakes exceed required levels. This may occur after acute or chronic exposure that results in internal concentrations that exceed the body's ability to regulate and excrete surplus metal ions. It is even possible that certain metal carcinogens, such as arsenic, may be essential. Interactions among essential and nonessential elements introduce complexity into the evaluation of metal toxicology. For example, selenium plays a role in alleviating toxicity due to mercury exposure and zinc can antagonize cadmium toxicity. Intake of zinc, sulfur, and iron are known to modulate copper deficiency or toxicity. Therefore, examination of the essentiality is critical for interpreting exposure and toxicity of metals.

16.7 TOXICOLOGY OF SPECIFIC METALS

Arsenic

Arsenic naturally occurs in the environment in several different forms. Inorganic arsenic (i.e., as arsenite (As^{3+}) or arsenate (As^{5+})) is the most common form of arsenic in soil, water, and air. On average, background concentrations of As in soil are approximately 10 mg/kg, while average concentrations in groundwater are less than 1 µg/l. Average concentrations in air away from a point source are generally under 5 ng/m^3 in rural area and 20–30 ng/m^3 in more urban settings (~20–30 ng/m^3 overall). For the typical individual, diet is the largest source of inorganic arsenic exposure, with intake levels in the U.S. population ranging from about 1 to 20 µg/day. Grains, fruits, and vegetables typically contain the highest levels of inorganic arsenic. Seafood, the largest source of total arsenic in the diet, typically contains high levels of arsenobetaine, a nontoxic organic arsenic compound.

Overall, human exposure to inorganic arsenic from these background sources is well below doses associated with adverse health effects. Several areas throughout the United States and the world, however, contain naturally occurring inorganic arsenic in drinking water, which constitutes a significant source of exposure (e.g., drinking water sources contain over several hundred micrograms per liter arsenic) and pose a public health concern. Organic arsenic compounds are also found in the environment. For example, arsenobetaine is a common constituent of seafood. The organic arsenic compounds monomethylarsonic acid (MMA) and dimethylarsinic acid (DMA) are also naturally occurring in environmental media, although usually at lower levels than inorganic arsenic. These two compounds are currently used as herbicides. Another form of arsenic with toxicological importance is arsine gas, which is formed when arsenic reacts with hydrogen gas. It is used in the semiconductor industry but also may be generated during the smelting and refining of metals, which involves treating the metals with acid. Arsine gas is a potent hemolytic agent (i.e., it breaks down blood cells).

When inorganic arsenic is ingested or inhaled, it undergoes sequential reduction and methylation reactions and is ultimately extensively metabolized to MMA and DMA, which are excreted in urine with a whole-body half-life of few days. Arsenic excreted in urine is typically composed of 10–30% inorganic arsenic, 10–20% MMA, and 60–80% DMA, although there is significant interindividual variability likely due to polymorphisms is arsenic-metabolizing genes. In contrast, when organic arsenic compounds are ingested, they undergo limited cellular uptake and are

TABLE 16.2 Health-Based Oral Intake Levels for Metals

		Recommended Dietary Allowance for Adults		Tolerable Upper Intake Level for Adults		Chronic Oral Reference Dose for Adults	
Element	Essential role	(mg/day)	(mg/kg-day)	(mg/day)	(mg/kg-day)	(mg/day)	(mg/kg-day)
Arsenic	Not essential for humans	—	—	—	—	0.02	0.0003
Boron	Not essential for humans	—	—	20	0.3	14	0.2
Calcium	Essential for teeth and bone formation	1000	14	2500	36	—	—
Chloride	Principal anion needed to regulate extracellular fluid volume and plasma osmolality	2300	33	3600	51	—	—
Chromium	Glucose and insulin metabolism	0.025–0.035	0.0004–0.0005	—	—	150 (Cr^{3+})	1.5 (Cr^{3+})
						0.2 (Cr^{6+})	0.003 (Cr^{6+})
Cobalt	Component of vitamin B_{12}, necessary for maintenance of hematological status	—	—	—	—	—	—
Copper	Component of metalloenzymes	0.9	0.01	10	0.14	—	—
Fluoride	Associated with calcified tissues and prevents dental caries	3–4	0.04–0.06	10	0.14	4.2	0.06
Iodine	Component of thyroid enzymes	0.15	0.002	1.1	0.02	—	—
Iron	Component of proteins (e.g., hemoglobin, myoglobin, and ferritin)	8–18	0.1–0.3	45	0.6	—	—
Magnesium	Cofactor for enzymes involved in glycolysis and oxidative phosphorylation	320–420	4.6–6.0	350	5	—	—
Manganese	Component of metalloenzymes important for bone formation and cholesterol, and carbohydrate metabolism	1.8–2.3	0.026–0.033	11	0.16	9.8	0.14
Molybdenum	Cofactor for enzymes involved in catabolism of sulfur amino acids	0.05	0.0006	2	0.03	0.35	0.005
Nickel	Not essential for humans	—	—	1	0.01	1.4	0.02
Phosphorous	Component of biological membranes and supports tissue growth	700	10	4	0.06	—	—
Potassium	Major intracellular cation necessary for normal function	4700	67	—	—	—	—
Selenium	Component of proteins important for defense against oxidative stress and regulation of other vitamins and hormones	0.06	0.0008	0.4	0.006	0.35	0.005

TABLE 16.2 Continued

Element	Essential role	Recommended Dietary Allowance for Adults		Tolerable Upper Intake Level for Adults		Chronic Oral Reference Dose for Adults	
		(mg/day)	(mg/kg-day)	(mg/day)	(mg/kg-day)	(mg/day)	(mg/kg-day)
Sodium	Principal cation needed to regulate extracellular fluid volume and plasma osmolality	1500	21	2300	33	—	—
Vanadium	Not essential for humans	—	—	1.8	0.03	—	—
Zinc	Component of or catalyst for a number of enzymes	8–11	0.11–0.16	40	0.57	21	0.3

1. "—" = none available.
2. Range for male and female adults.
3. Reference dose = An estimate (with uncertainty spanning perhaps an order of magnitude) of a daily oral exposure to the human population (including sensitive subgroups) that is likely to be without an appreciable risk of deleterious effects during a lifetime. It can be derived from an no observed adverse effect level (NOAEL), an LOAEL, or a benchmark dose, with uncertainty factors generally applied to reflect limitations of the data used. Generally used in US EPA's noncancer health assessments.
4. Recommended daily intake = A daily dietary intake sufficient to meet the nutrient requirements of nearly all (97–98%) healthy individuals.
5. Tolerable upper intake level is the highest level of daily nutrient intake that is likely to pose no risk of adverse health effects to almost all individuals in the general populations.
6. Dose adjustments made based on a 70-kg adult.

excreted in urine largely without further metabolism. A key exception is in the rat, where ingestion of DMA results in relatively high amounts of a trimethylated arsenic metabolite (trimethylarsine oxide). While the methylation of arsenic has traditionally been considered a detoxification mechanism because it facilitates excretion, research over the past decade has shown that the process of arsenic metabolism produces unstable trivalent methylated intermediates. It is hypothesized that these intermediates play a role in the chronic adverse effects associated with arsenic exposure.

The potent toxicity of arsenic has been well known for centuries; it has been used as a homicidal agent throughout history, purportedly in several high-profile political assassinations. Due to its toxic properties, it has also been used as a pest control agent. In particular, in the early 1900s, lead arsenate was used as a pesticide in apple orchards. Although the use of lead arsenic has been phased out, because of its environmental persistence, this legacy contamination still poses significant public health concerns in some areas of the United States. Some modern uses of arsenic-based pesticides still exist, including the selective use of copper chromated arsenate as wood treatment and the use of organic arsenicals as a herbicide for cotton.

Most information regarding acute arsenic toxicity in humans derived from its medicinal use or from incidents in which individuals were poisoned with arsenic (either intentionally or accidentally). Signs of acute arsenic toxicity include encephalopathy, nausea, anorexia, vomiting, abdominal pain and diarrhea, dermatitis, muscle cramps, cardiovascular effects, hepatotoxicity, bone marrow suppression and hematologic abnormalities (anemia), vascular lesions, and peripheral neuropathy (motor dysfunction and paresthesia). Arsenic toxicity can often effectively be treated with the chelator 2,3-dimercapto-1-propanesulfonate or meso-2,3-dimercaptosuccinic acid.

Information on the oral chronic effects of arsenic are mainly from observations where humans have been exposed to high concentrations of arsenic in drinking water. Heath effects from arsenic in soil have not been observed, likely because of the reduced bioavailability of arsenic in soils and because overall, exposures via soil pathways are relatively low. Based on several studies, an increase in arsenic body burden is not observed until arsenic concentrations in soil exceed several hundred micrograms per kilogram of soil.

Based on drinking water studies, arsenic is a known bladder, lung, and skin carcinogen via the oral route. An association between arsenic and liver and kidney cancer has also been suggested, but the evidence is not as strong. Overall, cancer effects are not observed until drinking water exposures exceed several hundred micrograms per liter. Several noncancer effects are also associated with chronic oral exposure to arsenic. One of the most common and characteristic effects of arsenic ingestion is a pattern of skin changes that include generalized hyperkeratosis and formation of hyperkeratotic warts or corns on the palms and soles. Other key noncancer effects that occur at high exposures include cardiovascular disease, diabetes, and respiratory deficits. Occupational studies, particularly from copper smelters, provide the best information on the health effects of arsenic via inhalation. By the inhalation route, there is strong evidence of an increased risk of lung cancer; noncancer effects from high level worker exposure can include respiratory irritation, nausea, and neurological symptoms. High-level exposures can cause perforation of the nasal septum.

Health effects from dermal exposure to arsenic are rare. Arsenic is not readily absorbed into the bloodstream through skin and has not been observed to pose a significant health risk. The effects that have been identified include local irritation and dermatitis.

In general, because of extensive human exposure and their unique sensitivity, animal studies are used less frequently than with other metals to characterize arsenic toxicity; although they are used extensively to understand mechanistic aspects of arsenic toxicology and to understand disease mode of action. Until recently, there was no animal model for inorganic arsenic carcinogenicity. New research has suggested, however, that transplacental exposure to mice during critical periods of development will increase tumor development in mice during adulthood.

Both US Environmental Protection Agency (US EPA) and International Agency for Research on Cancer (IARC) have classified arsenic as a known human carcinogen. For Integrated Risk Information System (IRIS), US EPA has established an inhalation unit risk (IUR) of 4.3×10^{-3} $(\mu g/m^3)^{-1}$. The oral slope factor for arsenic currently in IRIS of 1.5 $(mg/kg\text{-}day)^{-1}$ is based on skin cancer. More recent evaluations by US EPA and other agencies have used bladder and lung cancer as the end point of concern because of their greater public health significance. While all key assessments to date have assumed linear extrapolations to assess low-dose arsenic carcinogenicity even though arsenic is not considered to be mutagenic, there is substantial epidemiological evidence in support of a non-linear, threshold dose–response. Revisions to the current IRIS cancer assessment are underway and include an evaluation by the National Academy of Sciences (NAS) for noncancer assessments; the current RfD for arsenic is 3×10^{-4} mg/kg-day; a reference concentration (RfC) has not been developed under the IRIS program. The oral reference is also anticipated to undergo revision in the near future. In 2001, US EPA established a maximum contaminant level (MCL) for arsenic of 10 µg/l.

Health-protective occupational levels for inorganic arsenic exposure have been developed by Occupational Safety and Health Administration (OSHA)[1] (8-h time-weighted average (TWA) of 10 $\mu g/m^3$ for inorganic arsenic compounds and 0.5 mg/m^3 for organic arsenic compounds) and American Conference of Industrial Hygienists (ACGIH) (TWA threshold limit value (TLV®) of 0.01 mg/m^3). The 15-min ceiling limit for arsenic developed by National Institute for Occupational Safety and Health (NIOSH) is 2 $\mu g/m^3$.

Beryllium

Beryllium is a hard, grayish metal that occurs as a chemical component of certain rocks, coal and oil, soil, and volcanic dust. Most of the mined beryllium is converted into alloys, which are used in making electrical or electronic parts and construction materials. Beryllium alloys with high content (40–100%) are commonly used in electric connectors and relays, bearings, and optical systems in the aerospace industry, which is an example of an important occupational exposure pathway. Beryllium enters the environment (air, water, and soil) as a result of natural and human activities, and human exposures primarily occur in industrial settings. Since a major source of atmospheric beryllium is coal combustion, the most prevalent chemical form is probably beryllium oxide. In pH-neutral environments, beryllium chlorides, sulfates, nitrates, and fluorides are soluble, while oxides and hydroxides are insoluble.

In general, exposure to water-soluble beryllium compounds poses a greater threat to human health than does exposure to water-insoluble forms. Judging from animal studies, the several different forms of beryllium that have been studied are poorly absorbed through both the GI tract and the skin. The most important route by which beryllium compounds are taken up by animals and humans is inhalation. In humans, beryllium particles may reside in the lung for several years, depending on particle size and duration of exposure. Soluble beryllium may distribute to the liver, lymph nodes, spleen, heart, muscle, skin, and kidney. Unabsorbed beryllium is primarily excreted through the fecal route, while soluble forms are excreted via the urine. Soluble forms of beryllium may cross the placental barrier based on rodent studies.

There are very few dose–response data regarding the health effects of orally administered beryllium and beryllium-containing compounds in humans. Oral toxicity studies in animals indicate that the GI system and skeletal system are target organs. Beryllium rickets (defective bone growths) have been observed in animal models and may be the result of competitive binding of phosphorous. Reproduction and development has not been adequately studied via oral exposures.

Inhalation exposure of animals and humans to beryllium can result in two types of potentially fatal nonneoplastic respiratory disease: acute pneumonitis and chronic beryllium disease (berylliosis). Lethality and decreased longevity appear to be due to the development of chemical pneumonitis. A 1948 investigation of acute beryllium pneumonitis in three U.S. beryllium plants reported that all of the cases of beryllium pneumonitis studied were associated with inhalation exposures to beryllium concentrations of greater than 0.1 mg/m^3, primarily as beryllium sulfate or beryllium fluoride. For workers who were exposed after 1950, beryllium pneumonitis has been virtually eliminated except in cases of accidental exposure to concentrations above the OSHA standard of 0.002 mg/m^3.

Exposure to mean levels of 1 $\mu g/m^3$ for an extended duration may result in chronic beryllium disease, characterized by granulomas. Several retrospective cohort studies of workers who were exposed to beryllium from the 1940s to the 1970s report significantly higher mortality rates in comparison with the U.S. general mortality rates for the time periods studied.

[1] Occupational standards developed by OSHA are legally enforceable exposure limits, while values developed by ACGIH or NIOSH are recommended guidelines not enforced by regulatory agencies.

Most of the workers reportedly experienced shortness of breath, general weakness, and weight loss, and autopsies revealed granulomatous lung disease, lung fibrosis, and heart enlargement.

As in humans, animal studies indicate that the respiratory tract is the primary target for inhalation exposure to beryllium and some of its compounds. Pneumonitis, with accompanied thickening of the alveolar walls and inflammation of the lungs, was reported in rats and mice exposed to beryllium for 1 h at approximately 3.3 and 7.2 mg/m^3 (as beryllium sulfate), respectively, for 12 months or less. Some animal species exhibit hematological effects from beryllium exposure. Acute exposure had little hematological effect, but intermediate-duration exposure resulted in anemia in several animal species. Weight loss has been reported in some animal species after inhalation exposure to beryllium compounds.

No studies were discovered regarding death, systemic effects (other than some dermatological abnormalities), immunological effects, neurological effects, developmental effects, reproductive effects, genotoxic effects, or cancer in humans after oral or dermal exposure to beryllium or its compounds.

A number of studies have associated inhalation exposure to beryllium with an increased incidence of human lung cancer. In general, these studies have been judged to have limited application due to inadequate controls in the studies related to confounding factors such as smoking, improperly calculated expected deaths from lung cancer, as well as the use of inappropriate comparative control populations. Some beryllium compounds have been shown to be carcinogenic in animals. Rats exposed to concentrations of 0.035 mg/m^3 as beryllium sulfate for 180 days exhibited increased lung cancer rates, compared with controls. In another study, 18 of 19 rats exposed to beryllium concentrations of 0.62 mg/m^3 as beryl ore developed tumors that were classified as bronchial alveolar cell tumors, adenomas, adenocarcinomas, or epidermoid tumors. Based on limited human evidence and sufficient animal evidence, inhaled beryllium is considered a probable human carcinogen.

Dermal exposure may result in conjunctivitis and dermatitis of the skin. Beryllium is a skin sensitizer and can cause a delayed-type hypersensitivity reaction. Insoluble forms may accumulate in the skin and cause a chronic granulomatous lesion. Specific skin tests have been developed to identify beryllium-sensitive individuals.

The current daily oral RfD for beryllium is 2×10^{-3} mg/kg. The inhalation RfC is 2×10^{-2} μg/m^3. US EPA has classified beryllium as a B1 carcinogen (probable human carcinogen) by the inhalation route of exposure and has developed an IUR slope factor (2.4×10^{-3} per μg/m^3) for assessing cancer risks. ACGIH and OSHA also consider beryllium to be a potential carcinogen. The occupational exposure guidelines developed by ACGIH and OSHA TWA (8 h/day, 40 h/week) are 0.00005 mg/m^3 (TLV) and 0.002 mg/m^3 (permissible exposure limit; PEL), respectively. US EPA has set a drinking water MCL for beryllium at 0.004 mg/l.

Cadmium

Cadmium is usually not found in the environment as a pure metal; rather it is typically found as a mineral (such as cadmium oxide, chloride, sulfate, or in association with zinc). These solids may dissolve in water, and small particles of cadmium may be found in the air. Food and cigarette smoke are significant sources of cadmium exposure for the general public. The majority of cadmium is used for nickel–cadmium batteries and also for electroplating or galvanizing alloys due to is noncorrosive properties. Other uses include paint pigments, solders, plastic stabilizers, and specialty alloys.

The form of cadmium in food and water is generally the cadmium ion. Absorption of cadmium through the GI tract is limited (5–10% of the administered dose) and is dependent on the iron stores in the body, with low iron levels correlated with increased cadmium absorption. Absorption from the respiratory tract ranges from 5 to 35%, depending on the compound and particle size. Once cadmium enters the body, it is strongly retained in a number of organs. Cadmium absorbed by the human body is eliminated slowly (0.001% excreted per day), with the biological half-life estimated to be 10–30 years and is accumulated throughout a lifetime with over 30% of the body burden stored in the kidneys. Chelation therapy is typically not used in acute poisonings since these therapies can have significant adverse effects.

The kidney is the primary target organ of cadmium toxicity following chronic or subchronic exposure to cadmium via ingestion. The toxicity of cadmium to proximal renal tubular function is characterized by the presence of low- (but sometimes high-) molecular-weight proteins in the urine (termed proteinuria). Tubular dysfunction develops only after cadmium reaches a minimum threshold level in the renal cortex. Negative effects on calcium metabolism may occur as a result secondary to kidney damage.

Excessive oral exposure to high concentrations of cadmium causes severe irritation to the GI epithelium, resulting in nausea, vomiting, abdominal pain, and diarrhea. Painful bone disorders (Itai–Itai disease), which are characterized by osteomalacia and osteoporosis, have been observed in some humans chronically exposed to cadmium via diet. Decreased calcium content of bone and increased urinary calcium excretion are common findings in rats and mice following oral exposure to cadmium. Limited information is available to indicate cadmium as a neurotoxicant, and the blood–brain barrier limits cadmium exposure to the central nervous system (CNS). Human studies suggest a possible association between cadmium exposure and male reproductive toxicity, and animal studies suggest that exposure during pregnancy may result in fetotoxicity and developmental toxicity. Evidence from human and animal

studies do not indicate that cadmium is a carcinogen via the oral route.

Inhalation exposure to high levels of cadmium oxide fumes or dust can cause severe irritation to respiratory tissue. As for oral exposures, the kidney is also the main target organ for cadmium exposure via inhalation. Cadmium can induce obstructive lung disease that can progress from chronic bronchitis, fibrosis, pneumonitis, and pulmonary edema leading to emphysema. Lung injury following cadmium exposure can be at least partially reversible. There is persuasive evidence from studies in rats, but not mice, and equivocal results in hamsters, to conclude that chronic inhalation exposure to cadmium chloride is associated with increased frequency of lung tumors. Some epidemiologic studies in humans have suggested that inhaled cadmium is a pulmonary carcinogen. Smoking is a common confounding factor in the determination of the role of cadmium in lung cancer, and cadmium is an identified component of cigarette smoke.

Cadmium compounds have not been observed to cause significant health effects when exposure is by the dermal route. Similarly, minimal effects (e.g., lacrimation) have been observed following ocular exposure to cadmium.

The current daily oral RfD for cadmium is 5×10^{-4} mg/kg for water sources, in comparison with 1×10^{-3} mg/kg daily from food sources. US EPA has categorized cadmium as a class B1 carcinogen (probable human carcinogen) by inhalation only and has developed an IUR slope factor of 1.8×10^{-3} per μg/m^3. The ACGIH TWA-TLV® for cadmium is 0.01 mg/m^3 (0.002 mg/m^3 in the respirable fraction), and the OSHA PEL is 0.005 mg/m^3. US EPA has set a drinking water MCL for cadmium at 0.005 mg/l.

Chromium

Chromium is an element that primarily exists in the environment in two valence, or oxidation, states: trivalent chromium (Cr^{3+}) and hexavalent chromium (Cr^{6+}). Most naturally occurring chromium is in the most stable trivalent chromium form, while much of the hexavalent chromium in the environment is attributable to anthropogenic uses, such as applications in wood preservatives, pigments, and metal finishing. The primary route of exposure to chromium in the general population is food ingestion, with lesser potential exposures from inhaling ambient air and ingesting drinking water containing low levels of chromium. Workers in chromium-related industries can be exposed to concentrations of chromium elevated above ambient levels.

Human health risks posed by exposure to chromium are valence-specific. In general, hexavalent chromium compounds are more toxic than trivalent chromium compounds. Trivalent chromium is an essential nutrient that plays a role in glucose, fat, and protein metabolism by potentiating the action of insulin (Table 16.2). In contrast, hexavalent chromium is irritating and can cause adverse effects specific to, but not limited to, the portal of entry. Hexavalent chromium compounds are reduced to trivalent chromium in the presence of oxidizable organic matter, such as in the gastric environment and to a lesser extent in the lungs. Hexavalent chromium that escapes reduction is much more readily absorbed into cells than trivalent chromium. While reduction of hexavalent chromium to trivalent chromium outside of cells is a major protective mechanism, intracellular reduction of hexavalent chromium is thought to be an important mechanism of toxicity. Absorbed chromium is distributed to nearly all tissues and is predominantly excreted in urine.

Human data on the oral toxicity of hexavalent chromium are limited, except in cases of accidental or intentional ingestion of fatal or near-fatal doses. The few studies that involve chronic oral exposure suggest the potential for relatively high levels of chronic hexavalent chromium ingestion to cause GI, immunological, and hematological effects. The capacity for reduction of hexavalent chromium to trivalent chromium in GI fluids limits the potential for toxicity of ingested hexavalent chromium at plausible environmental concentrations. As in humans, there is evidence that high oral exposures to hexavalent chromium in animals can cause GI, immunological, and hematologic effects. A number of oral exposure animal studies have shown that hexavalent chromium can affect end points related to sperm or testes. Hexavalent chromium is a developmental toxicant, though these effects have been observed only at high doses associated with maternal toxicity.

There is no reliable evidence for a causal link between oral hexavalent chromium exposure and cancer in humans. Almost all of the studies evaluating the possible link between oral hexavalent chromium exposure and cancer effects come from a study population in China. The several studies published on this population show inconsistent results and are limited by the available exposure information. In the only reliable lifetime drinking water study in animals, the National Toxicology Program (NTP) found that rats developed carcinomas of the mouth and mice developed various types of tumors in the small intestine. The hexavalent chromium concentrations resulting in increased tumors were much higher than plausible environmental concentrations, and thus the findings are of questionable relevance to humans at any plausible exposure levels.

The respiratory tract is the major target of inhalation exposure to hexavalent chromium compounds in humans and animals. The preponderance of human data on noncancer effects of inhaled hexavalent chromium, including nasal mucosal irritation, atrophy, ulceration, and perforation, are from studies of workers exposed to chromic acid mists in the chrome plating industry. These likely represent worst-case exposures for noncancer effects of hexavalent chromium because of the acidic and oxidative nature of chromic acid. Chromic acid exposures are unlikely to occur outside of industrial settings, whereas hexavalent chromium particulate

is a more environmentally relevant form. Animal studies indicate that adverse effects from exposure to particulate hexavalent chromium start at higher concentrations and in different areas of the respiratory system, than those of dissolved hexavalent chromium compounds. Some nonrespiratory effects of inhaled hexavalent chromium, including renal and GI effects, have been reported in occupational studies. These effects generally occur at higher concentrations than those for the most sensitive respiratory effects.

Inhalation of hexavalent chromium at sufficiently high concentrations is associated with respiratory system cancers in workers. This is based on occupational epidemiology studies, most of which evaluated workers who were exposed to very high hexavalent chromium concentrations compared with current workplace exposures. There is a good correlation between the dose of chromium, expressed as a function of concentration and time of exposure, and the relative risk of developing lung cancer. The best dose–response relationships have been developed from studies of workers in the chromate production industry. While few available studies have examined cancer mortality in relation to nonoccupational populations exposed to environmental hexavalent chromium (i.e., residing near industrial sites that produced or used chromium products), they have not found increases in lung or total cancer mortality.

Dermal effects have been reported in some occupational settings where there is potential for splashing, spilling, or other skin contact with hexavalent chromium–containing material, such as might occur during chrome plating or work with wet cement. The dermal effects of hexavalent chromium are the result of two distinct processes: (1) irritant reactions, such as skin ulcers and irritant contact dermatitis; and (2) delayed hypersensitivity (allergic) reactions, that is, allergic contact dermatitis (ACD), in chromium-sensitized individuals. Symptoms are similar for both, and intensity of the reaction is a function of dose in both cases. At sufficiently high exposures, hexavalent chromium can cause ulcers or "chrome holes." Dermal exposure to chromium compounds is not known to result in cancer in humans or animals.

The US EPA oral RfDs for hexavalent chromium and trivalent chromium are 0.003 and 1.5 mg/kg-day, respectively. US EPA has developed separate inhalation RfCs for hexavalent chromium. For chromic acid mists and dissolved hexavalent chromium aerosols, the RfC is 8×10^{-6} mg/m^3. For hexavalent chromium particulates, the RfC is 1×10^{-4} mg/m^3. There is no inhalation RfC for trivalent chromium. US EPA has classified hexavalent chromium as a known human carcinogen by the inhalation route of exposure but has not determined its oral carcinogenicity. Trivalent chromium is not classified by US EPA as a carcinogen by any route of exposure. The toxicity and carcinogenicity of hexavalent chromium is currently undergoing reevaluation by US EPA. The US EPA MCL for drinking water is 0.1 mg/l and is for total chromium. The OSHA 8-h TWA PELs for hexavalent chromium and trivalent chromium are 0.005 and 0.5 mg/m^3, respectively. ACGIH has published two TLVs for hexavalent chromium: 0.05 mg/m^3 for water-soluble hexavalent chromium and 0.01 mg/m^3 for insoluble hexavalent chromium.

Copper

Naturally occurring copper is found in various ores, primarily chalcopyrite ($CuFeS_2$), cuprite (Cu_2O), and malachite (Cu_2)(CO_3)$(OH)_2$. Copper exists in the environment in three major valence states: metallic (Cu^0), cuprous (Cu^{1+}), and cupric (Cu^{2+}). Under aerobic conditions, copper in soil and water is predominantly present in the cupric form, where it tightly binds to organic matter. Complexes with the cupric form of copper are very colorful; for example, copper aluminum hydroxyphosphate provides the bluish-green color to the turquoise gemstone. Copper can be released into the environment due to mining and smelting operations, incineration, industrial discharges, and weathering of soil. Copper can leach into drinking water from copper-containing pipes and plumbing fixtures.

Copper is also an essential nutrient required for a variety of biological processes, including growth, cardiovascular function, neuroendocrine function, and iron metabolism (Table 16.2). Copper is incorporated into enzymes involved in aerobic metabolism (e.g., mitochondrial cytochrome c oxidase) and is a cofactor in copper–zinc superoxide dismutase, which protects cells from free radical damage.

The majority of ingested copper is absorbed in the small intestine (the duodenum and ileum), with smaller amounts absorbed in the stomach. The fraction of copper absorbed can vary significantly based on the form of copper and the amount of copper in the diet, ranging from approximately 10% for diets high in copper to as high as 70% for diets low in copper. Excess zinc, cadmium, and molybdenum can reduce copper absorption, metabolism, and retention. Ascorbic acid, fiber, fructose, and other dietary components can also interfere with copper absorption. Absorption is also age-related, with greater absorption during the neonatal period.

Copper that is absorbed likely enters intestinal cells via passive diffusion, possibly via high-affinity membrane permeases. Copper is then transported through the basolateral membrane via specific transporter proteins. Copper is distributed in the body bound primarily to albumin and amino acids, with some copper also bound to ceruloplasmin. Following absorption, copper is rapidly secreted into the bile, resulting in net absorption of 10–20% of ingested copper in adults. Infants retain a greater percentage (20–30%) due to increased demand related to tissue growth and increased expression of copper-containing proteins. Small amounts of copper are also excreted in the hair and sweat.

Because copper is an essential nutrient, several homeostatic mechanisms prevent the absorption of excess copper

once physiologic needs have been met. The first barrier to prevent absorption of excess copper is the GI lining, where excess copper binds to metallothionein in GI mucosal cells and is excreted when these cells turn over. As a second mechanism, any excess copper that passes through the GI lining is stored in the liver or incorporated into bile and excreted. Homeostatic mechanisms for regulating GI absorption and clearing copper from the body are not fully developed in children less than 1 year old.

There are a few rare genetic diseases that affect copper absorption and metabolism. Menkes syndrome is a congenital disorder characterized by a defect in intestinal absorption of copper. Individuals with Wilson's disease have a genetic defect that affects copper metabolism, which results in an excessive accumulation of copper in the liver and brain.

In contrast to most chemicals, for which the chronic health effects typically occur at lower doses or exposure concentrations than the acute effects, the toxicity of copper is unusual in that acute effects following oral exposure can, under some circumstances, occur at lower exposure levels than chronic effects. Acute effects of copper ingestion consist predominantly of GI symptoms, including nausea, abdominal pain, and vomiting. These effects are probably related to the presence of free Cu^{2+} ions in the stomach. The GI effects of copper generally occur directly following exposure and are readily reversible once exposure ceases. Moreover, at levels of copper intake that cause GI symptoms, there is no evidence of systemic copper toxicity, such as effects on the liver or kidney. This is in part because the acute toxic response to copper occurs prior to its absorption and distribution throughout the body. There is some evidence that children may be more susceptible to acute copper effects than adults.

The liver is the primary target organ for chronic exposure to copper, as demonstrated by studies in laboratory animals that indicate liver effects typically occur at lower doses than effects in other organ systems, such as the kidney, hematopoietic, cardiovascular, or central nervous systems. Liver effects observed in rats following exposures of 13–15 weeks include increased activity of liver enzymes, inflammation, regeneration of parenchymal tissue, and chronic hepatitis. There is little evidence of liver toxicity in humans due to chronic exposure to copper, except for individuals with preexisting liver diseases (such as Wilson's disease) who are unusually susceptible to copper.

There is no evidence that copper causes cancer in either humans or animals via oral exposure. There are no studies or reports of cancer in humans following oral exposure to copper. Increases in cancer were not observed in animals exposed to copper via oral exposure. These studies, however, are considered inadequate for evaluating potential carcinogenicity of copper due to study design deficiencies.

Copper is considered to be a respiratory irritant. Effects in workers attributed to chronic copper dust exposure, at concentrations ranging from 111 to 464 mg/m^3, include neurological changes, digestive disorders, and pulmonary complications. Copper has also been associated with "vineyard sprayer's lung," which involves formation of pulmonary granulomas with copper inclusions, and effects on alveolar macrophages, and which has been reported for workers exposed to anti-mildew agents containing from 1 to 2.5% copper sulfate. Pulmonary effects have also been observed in animals following inhalation exposure to copper, including decreased beating of cilia in hamsters exposed to 3.3 mg/m^3 copper (as copper sulfate) for 3 h and alveolar thickening in mice exposed to 0.12 mg/m^3 copper (as copper sulfate) for 3 h/day, 5 days/week, for 1–2 weeks. There are no studies or reports of cancer in humans or animals following inhalation exposure to copper.

There is limited evidence that dermal exposure to copper can result in ACD in some people. At a turbine generator–manufacturing facility, workers exposed to copper dust developed green skin discoloration but no evidence of systemic copper absorption.

US EPA has not developed an RfD for their Agency for Toxic Substances and Disease Registry (IRIS) database. ATSDR's oral and acute minimal risk level (MRL) for copper is 0.01 mg/kg-day, based on GI effects in humans. The MCL for copper is 1.3 mg/l, as an action level, and is based on GI effects in humans. OSHA has developed a TWA-PEL of 1 mg/m^3 copper for copper as dusts and mists and a TWA-PEL of 0.1 mg/m^3 for copper as a fume, while the ACGIH TWA-TLV is 0.2 mg/m^3 for copper fumes. US EPA has assigned copper a Weight-of-Evidence Classification for human carcinogenicity of "D" because there are no human data, inadequate data from animal cancer assays, and equivocal mutagenicity data.

Lead

Lead is a naturally occurring bluish-gray metal found in the earth's crust, in ore deposits that are readily accessible and widely distributed throughout the world. Lead exists in three oxidations states: Pb^0, Pb^{2+}, and Pb^{4+}, but it exists in nature primarily in the Pb^{2+} state, in a variety of lead compounds. The majority of organolead compounds are in the Pb^{4+} state. Most of the lead mobilized in the environment is the result of human activities.

Absorption of lead from the GI tract is thought to occur via passive diffusion processes—both through and between intestinal cells, as well as via specialized energy-dependent transporters that are involved in absorption of calcium and iron. The primary factors that affect lead absorption include age and dietary factors. Lead absorption is greater in children than adults. Absorption is also greater when ingested on an empty stomach, than when there is food in the gut. Both iron and calcium in the diet can reduce lead absorption.

Following absorption, lead is either distributed to various tissues in the body or excreted. Within the body, lead

is primarily distributed to the bone, which accounts for approximately 73% of the total lead body burden in children and approximately 94% in adults. Bone lead can also contribute to blood lead, with 40–70% of lead in the blood derived from the bone. During childhood, bone lead is readily transferred back to the blood. Lead is excreted in both the feces and urine, with fecal excretion greater than urinary excretion, and with fecal excretion increasing with increasing lead intake. The rate at which lead is cleared from the blood (due to both distribution to other tissues and excretion) can vary depending on prior exposure to lead.

Lead is also absorbed from the respiratory tract following inhalation. Dermal absorption of inorganic lead compounds is generally considered to be much less than absorption by oral or inhalation routes of exposure.

The toxic effects of lead are the same regardless of the route of entry into the body. Exposure to lead is typically based on the amount of lead in the blood, in microgram lead per deciliter of blood (μg/dl), referred to as the BLL. Very high levels of lead intake can cause kidney damage, convulsions, coma, and even death. Reversible proximal tubular damage of the kidneys can result from acute lead exposure. Heavy, chronic exposure can cause nephritis, interstitial fibrosis, and tubular atrophy. Colic—characterized by abdominal pain, constipation, cramps, nausea, vomiting, anorexia, and weight loss—is a symptom of lead poisoning in occupationally exposed cases and in children. Lead can also disrupt heme biosynthesis. Lead affects the hematopoietic system by altering the activity of three enzymes involved in heme biosynthesis. The impairment of heme synthesis has a number of subsequent effects, including decreased hemoglobin levels and anemia. These effects have been observed in lead workers and in children with prolonged lead exposure. A decrease in cytochrome P450 content of hepatic microsomes has been noted in animal studies.

At lower levels of exposure, lead can cause neurological, cognitive, and behavioral effects. In adults, slowing of nerve conduction velocity has been observed at BLLs above 20–30 μg/dl. Other subclinical effects have been observed at higher BLLs (≥30 μg/dl), typically in subjects who were chronically exposed. Effects include impaired performance on cognitive, verbal reasoning, and memory tasks; mood disturbance; lethargy; fatigue; dizziness; impaired visual–motor coordination; and postural balance.

The most sensitive target associated with chronic, low-level lead exposure is the developing nervous system in young children. Over the past several decades, many studies have reported declines in IQ associated with increasing BLL. The effect of lead on IQ is subtle; some investigators have estimated that for each increase in blood lead of 10 μg/dl, there is a decrease in average IQ scores of 1–5 points. While effects of lead on IQ may have significant implications for the population as a whole, the health implications of low-level lead exposure are difficult to discern on an individual basis. More recently, researchers have also reported behavioral effects, such as increased distractibility and decreased attention span, associated with increasing BLLs.

Several studies of workers exposed to high levels of lead have reported increased cancer risks, most commonly for stomach and lung cancer. However, these studies were limited by lack of adjustment for other exposures, such as arsenic, or other factors associated with these cancers, including smoking and dietary factors. Renal tumors have been observed in animals following exposure to lead.

US EPA has not established any toxicity criteria for lead. Instead, lead risks are evaluated by modeling blood lead concentrations based on maintaining BLLs below 10 μg/dl. Due to concerns for lead at lower BLLs, recently, the Centers for Disease Control established 5 μg/dl as a lead reference level to protect against adverse health effects. For assessing potential health risks and establishing environmental cleanup criteria, BLLs in children are typically modeled using US EPA's Integrated Exposure Uptake Biokinetic model; BLLs in adults are typically modeled using US EPA's Adult Lead Model. The MCL for lead is 0.015 mg/l, as an action level, with a maximum contaminant level goal (MCLG) of zero. OSHA has developed several occupational limits for lead compounds. OSHA has set an 8-h airborne action level of 30 $\mu g/m^3$ and above this level employees must wear a respirator. OSHA's TWA-PEL and ACGIH TWA-TLV for inorganic lead compounds (including lead phosphate) are both 0.05 mg/m^3, as lead. However, if an employee is exposed to lead above the PEL for more than 30 days/year, engineering controls must be implemented to reduce exposure to 200 mg/m^3, and any combination of engineering, work practice, and respiratory controls may be implemented to reduce and maintain lead exposure to or below 50 $\mu g/m^3$.

US EPA has assigned lead a Weight-of-Evidence Classification for human carcinogenicity of "B2," a "probable human carcinogen," based on sufficient animal evidence but inadequate human evidence. Even though the weight of evidence for lead carcinogenicity is B2, US EPA does not evaluate lead cancer risks using a CSF, having concluded that neurological effects in young children are the most relevant end point.

Manganese

Manganese (Mn) is a naturally occurring element found in the environment in rocks, soil, water, and food. Manganese is an essential nutrient in humans and animals, playing a role in many biological functions, including metabolic regulation, cellular protection against oxidative stress, and bone mineralization and as a constituent in metalloenzymes (such as manganese superoxide dismutase) and as a cofactor for a number of enzymatic reactions (Table 16.2). Since manganese is required by the body, it is found in most diets

and is present in green leafy vegetables, nuts, grains, legumes, and some fruits. Manganese can also be found in drinking water but intake is typically much lower than from the diet. Manganese is also present in the gasoline additive methylcyclopentadienyl manganese tricarbonyl. Manganese intake from air is typically considered negligible but may be high in certain occupational environments. Although manganese is an essential nutrient, high exposures via inhalation or ingestion may result in adverse effect(s).

Humans generally maintain stable tissue levels of manganese from dietary intake through homeostatic regulation of GI absorption and hepatobiliary excretion. GI absorption of manganese is rapid and tends to be greater from more soluble forms of manganese, with solubility of manganese species generally following soluble sulfates > less soluble phosphates > less soluble oxides. Manganese absorption is higher in neonates compared with other age groups, likely due to an increased need for manganese during this life stage. A key determinant of manganese absorption is dietary levels of iron, with low iron leading to increased manganese absorption. Once absorbed into the circulation, manganese is transported to all tissues, with higher concentrations typically found in the liver, kidney, pancreas, and brain (striatum, globus pallidus, and substantia nigra).

Following inhalation exposure, manganese can be transported through the olfactory nerves in the nasal mucosa directly to the brain; however, the majority of respirable manganese is deposited in the lung, absorbed into the circulation system, and transported to all tissues. Recent PBPK models indicate that low levels of inhaled manganese are under the same homeostatic regulation as ingested manganese once absorbed into the circulation. Absorption in the lung is typically higher for more soluble forms of manganese. Similar to the ingestion pathway, absorption in the lung is also influenced by iron status, with iron deficiency resulting in enhanced absorption.

Manganese is not carcinogenic in animals or humans from ingestion or inhalation exposure but is a neurotoxin at high exposure concentrations. Although other systemic adverse effects have been observed at very high manganese exposure concentrations, neurological effects are the most sensitive end point in animals and humans. Suboptimal iron intake and abnormal liver function have been shown to result in increased concentrations of manganese in the brain and also increased susceptibility to manganese neurotoxicity.

There is conclusive evidence, based on epidemiological studies of occupationally exposed groups of individuals (such as welders, ferroalloy refinery workers, and manganese miners), that chronic inhalation of high concentrations of manganese (>1 mg/m^3) leads to adverse neurological effects in humans. A disabling neurological disease known as "manganism," the symptoms of which are altered gait and fine tremor, is associated with chronic inhalation of high concentrations of manganese. Some of the symptoms of manganism are similar to the symptoms of Parkinson's disease, so manganism is often referred to as "Parkinson's-like disease." Subclinical neurological effects (predominantly subtle neuromotor effects) have been observed from chronic low-level occupational manganese inhalation exposures (0.06 to <1 mg/m^3). Although potential neurological effects from very low levels of manganese in air (<0.01 mg/m^3) have been investigated in communities near manganese-emitting industries, the results of these studies have not been conclusive due to potential confounding factors that limit interpretation of results. Recent PBPK modeling of manganese tissue concentrations suggests that homeostatic regulation prevents accumulation of manganese in the brain at inhalation concentrations of less than 0.01 mg/m^3. Further, the PBPK modeling for all age groups suggests that fetuses, nursing neonates, children, and the elderly are not more sensitive than young adults to typical environmental levels of manganese in air.

Although there is sufficient evidence for adverse neurological effects in humans from inhalation exposure to high levels of manganese, the evidence that oral exposure leads to adverse neurological effects in humans is more limited. Adverse neurological effects similar to manganism have been reported in people consuming very high levels of manganese in drinking water (>10 mg/l). Several studies have reported subclinical neurological effects in children (neuromotor, behavioral, and cognitive) from exposure to manganese-contaminated drinking water (typically ranging from 0.6 to 1 mg/l). Limitations in these studies, however, such as cross-sectional study design, sample size, insufficient exposure estimates to identify all manganese sources, lack of control for coexposures to other neurotoxins, and lack of control for nutritional status such as dietary iron, make it difficult to draw any conclusions about a cause and effect relationship.

US EPA has developed an RfD for manganese based on CNS effects (0.14 mg/kg-day). The RfC (0.05 µg/m^3) was developed by US EPA based on occupational exposure resulting in impaired neurobehavioral function. The secondary MCL for manganese is 0.05 mg/l; it is not a health-based standard, rather it is based on aesthetic effects (e.g., taste, odor, or color). OSHA has developed a PEL (ceiling) of 5 mg/m^3 and the ACGIH TLV is now 0.2 mg/m^3 for inorganic manganese compounds. US EPA has determined that insufficient data are available to classify manganese as a carcinogen ("D" classification).

Molybdenum

Molybdenum (Mo) is a naturally occurring transition metal that can be found in the environment in several different valence states; the most common valence state for naturally occurring molybdenum is Mo^{4+}. In this form, molybdenum is usually complexed with sulfur to form the compound molybdenite (MoS_2). Not only is molybdenite the most

abundant form of molybdenum in ores but is also the most commercially valuable form. Other common molybdenite minerals include a lead complex called wulfenite ($PbMoO_4$) and a calcium complex called powellite ($CaMoO_4$). In soil, molybdenum is generally found adsorbed to iron or aluminum oxides, clay, and/or organic matter. In water, the molybdate ion (MoO_4^{2-}) dominates molybdenum aqueous speciation except under low pH (<4) and anoxic conditions.

Molybdenum is essential to normal biological function. Molybdenum serves as a cofactor for several enzymes in humans and animals that are important for metabolism of sulfur amino acid and heterocyclic compounds. Based on potential health concerns associated with molybdenum deficiency, the Institute of Medicine (IOM) of the NAS has developed recommended dietary allowances (RDAs). Consistent with these RDAs, recent studies of molybdenum metabolism have demonstrated that an intake of 43 µg/day would be sufficient to maintain plasma molybdenum levels at the necessary steady state in healthy adults (Table 16.2).

When ingested, water-soluble forms of molybdenum are readily absorbed by the GI tract, while poorly soluble compounds (e.g., molybdenum disulfide) are minimally absorbed. Overall, several key studies conducted in humans indicate that molybdenum absorption ranges from 28 to 95% following oral intake. Once absorbed, molybdenum distributes rapidly to the blood and most organs. Blood molybdenum concentrations have been reported to be on average 5 µg/l, but levels as high as 400 µg/l have been measured after elevated exposures. Upon exposure, the highest concentrations of molybdenum have been found in the kidney and liver, with lower levels in the adrenal glands and long bones. Molybdenum does not bioaccumulate in tissues and, after exposure cessation, tissue concentrations decrease to steady-state levels in a relatively short time frame in most organs. Molybdenum is excreted primarily via the urine or feces. Animal and human studies show similar excretion profiles and indicate that very little molybdenum is excreted via the bile. Molybdenum compounds have been found to readily cross the placental barrier.

Studies of chronic oral exposure to molybdenum in human populations have found a correlation between serum uric acid levels and several other biochemical end points (e.g., serum ceruloplasmin, a copper-carrying protein) with a gout-like sickness. This condition was characterized by pain, swelling, inflammation and deformities of the joints, and, in all cases, an increase in the uric acid content of the blood. In a number of cases, this condition was accompanied by illnesses of the GI tract, liver, and kidneys. Both serum molybdenum and serum xanthine oxidase activity were positively correlated with serum uric acid levels. Increased urinary excretion of copper was inversely correlated to increased serum levels of molybdenum.

Acute health effects observed in animal studies include diarrhea, coma, and death from cardiac failure. In animals, subchronic or chronic exposures mainly led to growth retardation; anemia; diarrhea; and changes to the thyroid, kidney, and liver. Molybdenum was also found to disturb bone metabolism, giving rise to lameness, bone joint abnormalities, osteoporosis, and high serum phosphatase levels. Elevated molybdenum exposure also was found to adversely affect reproduction (e.g., decreased gestation weight and offspring survival). The most common expression of molybdenum toxicity is a condition that occurs primarily in ruminants (e.g., cattle) called molybdenosis. This is a condition caused by a Mo-induced copper deficiency, is characterized by severe diarrhea and emaciation, and is usually fatal.

In occupational settings, there have been reports that inhalation of molybdenum (i.e., metallic molybdenum dusts or molybdenum trioxide (MoO_3)) may adversely affect health. Pneumoconiosis (restrictive lung disease) has been reported following inhalation exposure. Studies indicate that airborne concentrations ranging from 1 to 19 mg Mo/m^3 can induce signs of pneumoconiosis. Gout and multiple sclerosis have also been reported in several case studies of humans exposed to high molybdenum concentrations in air. Studies regarding the carcinogenicity of molybdenum compounds in humans are limited, and these studies suffer from poor study design and conflicting results.

Carcinogenicity studies in animals are also limited. Two-year inhalation studies (6 h/day, 5 days/week, 105 weeks) were conducted with rats and mice exposed to 0, 10, 30, or 100 mg/m^3 molybdenum trioxide. Negative (or equivocal) results were reported for rats, while limited evidence was reported for mice based on increased incidence of alveolar/bronchiolar adenomas or carcinomas. Molybdenum trioxide was also found to be weakly carcinogenic in a short-term lung adenoma assay with mice. Overall, there is some evidence for carcinogenic effects of molybdenum via inhalation in rodents; however, the human evidence is weak.

Limited studies are available describing dermal exposures to molybdenum compounds. Available information suggests some molybdenum compounds produce varying degrees of skin and eye irritation. For example, molybdenum trioxide and sodium molybdate (Na_2MoO_4) are strong eye and skin irritants; however, others, such as calcium and zinc molybdates are not primary irritants.

US EPA has developed an oral RfD of 0.005 mg/kg-day, which is based on a dietary intake that was associated with serum uric acid levels and a gout-like sickness in adults. This dose also satisfies the nutrient requirements for all healthy members of the population. US EPA determined that information is not sufficient to develop an inhalation RfC or to quantify potential cancer risks from exposure to molybdenum. Occupational OSHA standards (TWA-PELs) for insoluble and soluble manganese compounds are 15 mg/m^3 (as total dust) and 5 mg/m^3, respectively, while ACGIH has developed exposure guidelines for insoluble (ACGIH TLV®-TWA 3 mg/m^3 (respirable fraction)) and soluble

(ACGIH TLV-TWA 0.5 mg/m^3 (respirable fraction)) molybdenum compounds to protect against lower respiratory tract irritation. There is currently not a drinking water standard (MCL) for molybdenum; however, a nonenforceable health advisory level has been published by US EPA at 0.08 mg/l for 1- or 10-day exposure for a 10-kg child. In addition, using the RfD, EPA estimated a drinking water equivalent level of 0.2 mg/l based on a drinking water intake of 2 l/day and body weight of 70 kg/day for adults.

Mercury

Mercury is a ubiquitous part of man's environment arising from both natural and anthropogenic sources. It exists in three distinct forms: elemental mercury, inorganic mercury salts, and organic mercurials including both alkylated (e.g., methyl, dimethyl, and ethyl) and aromatic (e.g., phenyl) species. As with many metals, the chemical form of mercury is critical to understanding the movement of mercury in the environment, the potential for human exposure, and the possible human health effects.

Elemental mercury exists as a liquid at room temperatures and may be found in electronic devices such as switches, pressure gauges, and thermostats, particularly in older equipment. Elemental mercury is also found in small quantities in fluorescent lights, including compact fluorescent bulbs. Along with some ionic mercury, elemental mercury is present in coal combustion emissions, which represent the largest source of anthropogenic mercury contribution to the environment in the United States. As it is a major component in dental amalgam, exposures to elemental mercury have been of concern for dental workers, although the increasing use of composite fillings can be expected to result in decreases in exposure. Most studies have affirmed the safety of dental amalgam fillings for the general public. Due to its relatively high vapor pressure, spills of elemental mercury can volatilize and represent an inhalation exposure hazard. Elemental mercury is also well absorbed in the lungs, although its GI absorption is quite low.

Mercury salts (mercurous, Hg^+; mercuric, Hg^{2+}) are used as industrial catalysts, battery components, pigments, and preservatives, although the use of mercury compounds for many of these purposes has declined as other non-mercury–employing technologies have become available. Due to their low volatility and low lipid solubility, these forms present a lesser exposure hazard compared with elemental mercury. Of the two, mercuric salts typically have greater water solubility and better absorption in the GI tract.

Some organic forms of mercury are used as fungicides (e.g., phenylmercury acetate) and due to their highly lipophilic nature, they may present a dermal exposure hazard although they generally have low volatility and thus limited potential for inhalation exposure. The most notorious forms of organic mercury, methyl mercury and dimethyl mercury, are not produced industrially but arise from methylation of inorganic mercury by biota in aquatic environments. Due to their high lipid solubility, they bioaccumulate in the aquatic food chain and may reach high concentrations in top-level predator fish. The use of ethyl mercury as a medical preservative (as thimerosal or ethylmercury salicylate) has been substantially curtailed in most Western nations as a precautionary measure although the majority of studies in exposed human populations have not indicated any adverse effects.

As noted above, the form of mercury is highly important in determining potential health effects, in particular the target organ systems. Elemental mercury primarily targets the CNS and the kidney. The key symptoms associated with chronic elemental mercury exposure primarily involve tremor and other peripheral neuromuscular effects, changes in mood and mental state (e.g., excitability, irritability, and depression), and loss of cognitive ability. Kidney toxicity occurs after prolonged exposure as mercury accumulates in the proximal tubules leading to progressive tubular damage. The toxic sequelae of exposure to organic mercurials occur in the CNS, and in particular the CNS of developing children. The particular sensitivity of the developing nervous system was identified in high-level exposure cases in Japan and Iraq and more recently in several populations with high levels of fish consumption but no particular sources of mercury pollution. It should be emphasized that the neuropsychological test deficits observed in these studies were only associated with mercury at the population level and cannot be linked to mercury on an individual basis. Interpretation of the more recent epidemiology data is also complicated by coexposures to other chemicals associated with the subjects' diets, including polychlorinated biphenyls (PCBs), selenium, and omega-3 fatty acids. As with lead, a threshold for the effect of methylmercury on CNS development is yet to be defined. Epidemiology studies have also suggested that methylmercury exposure from high fish consumption is associated with adverse cardiovascular outcomes in adults (primarily myocardial infarction risk), although interpretation of the studies is complicated by the cardioprotective effects of omega-3 fatty acid consumption. Overexposure to the mercurous and mercuric salts are most typically characterized by a corrosive effect on the tissues, resulting in GI irritation, pain, and ulceration. As with elemental mercury, the mercury salts may also produce kidney damage after chronic exposure. Studies, chiefly conducted in animals or in human blood cells *in vitro*, have also suggested mercuric chloride may alter immune function.

The primary biomarkers for mercury exposure (regardless of form) are urinary mercury and blood mercury, the former being easiest to obtain but the latter being more relevant for assessing exposure, particularly for *in utero* exposures. Hair mercury is also frequently measured (e.g., such measurements form part of the basis for US EPA's methylmercury RfD). Data from the Centers for Disease Control (CDC) for

the 2003–2008 period indicate that average blood mercury level in women aged 18–49 was 0.7 μg/l during that period while that in children was 0.3 μg/l.

Due to the extensive database of human studies, animal studies have played a limited role in assessing the risks of mercury exposure. A large number of animal and *in vitro* studies have been employed in order to understand the potential mechanisms of action involved in mercury toxicity. This research has suggested a range of possible mechanisms (e.g., inactivation of various proteins via binding to sulfhydryl groups, induction of oxidative stress, and/or perturbation of intracellular calcium), many of which overlap or are interrelated. It is likely that no one mechanism predominates, a fact that further complicates study design and interpretation.

US EPA has established an RfC for elemental mercury (3×10^{-4} mg/m^3), an RfD for mercuric chloride (3×10^{-4} mg/kg-day), and an RfD for methylmercury (1×10^{-4} mg/kg-day). The US FDA has set a limit of 1 ppm for total mercury in all commercially caught fish. The analogous limits in Canada and the European Union are 0.5 ppm in both jurisdictions, with a limit of 1 ppm allowed for high-level predatory fish such as shark and swordfish. In developing these limits, regulators have struggled with the need to craft a risk communication message that balances the need to minimize mercury exposures in pregnant women (the population of concern) and at the same time encourages consumption of fish as a source of omega-3 fatty acids necessary for child neurological development. Occupational exposure levels for inorganic mercury compounds (as Hg) are 0.1 mg/m^3 (OSHA TWE-PEL) and 0.025 mg/m^3 (ACGIH TWA-TLV).

Nickel

Nickel is a hard, silver–white, metallic element used extensively in alloys and for plating because of its oxidation resistance. Nickel can exist in five valence states (Ni^{1-}, Ni^0, Ni^{2+}, Ni^{3+}, and Ni^{4+}), with the Ni^{2+} valence state being the most common under normal environmental conditions. Many nickel compounds can be formed from the Ni^{2+} ion, including organic, water-insoluble nickel compounds as well as inorganic nickel compounds that are either water-soluble or insoluble. Nickel, combined with other elements, occurs naturally in the earth's crust. Releases of nickel into the atmosphere occur from natural discharges, such as volcanic eruptions and windblown dust, and anthropogenic activities, such as fossil fuel combustion, nickel smelting and refining, and the production of steel or other nickel alloys. The four main categories of nickel compounds encountered in nickel production and refining processes, in decreasing order of solubility, are water-soluble nickel, sulfidic nickel, metallic (elemental) nickel, and oxidic nickel. In addition, nickel carbonyl, an organic and highly volatile liquid, is used industrially in the purification of metallic nickel and as a catalyst in chemical syntheses.

Nickel released to the atmosphere typically exists in particulate form or is adsorbed to particulate matter. Primary removal mechanisms of atmospheric nickel include gravitational settling and precipitation. Average nickel concentrations in ambient air typically range from 6 to 20 ng/m^3 and can be as high as 150 ng/m^3 near anthropogenic sources. Nickel released to soil may be adsorbed to soil surfaces depending on the soil conditions. The typical concentration of nickel in soil has been reported in the range of 4–80 ppm. Nickel released to aquatic systems generally exists in particulate forms that settle out in areas of active sedimentation, but nickel salts exhibit significant solubility in water. The nickel concentration in uncontaminated freshwater and seawater is low, generally around 0.3 μg/l.

Nickel is found in drinking water at an average concentration of about 2 μg/l. Assuming that a person consumes 2 l of water per day, the daily intake of nickel from drinking water is 4 μg/day. Many foods also contain nickel, and several dietary studies have reported average daily dietary intakes of nickel in the range of 69–162 μg/day. Nickel is an essential trace element in animals but has not been shown to be essential in humans (Table 16.2).

The clearance of nickel compounds from the respiratory system after inhalation exposure can range from a few days to a few months, with water-soluble nickel compounds exhibiting faster clearance than less-soluble forms. Absorption studies indicate that approximately 20–35% of inhaled nickel that is retained in the lungs is absorbed into the bloodstream. Nickel absorption following oral exposure varies from 3 to 40%, with greater absorption from drinking water than from consuming food. Although nickel is able to penetrate the skin, the majority of it remains in the skin layers rather than entering the bloodstream. Once absorbed into the bloodstream, nickel is excreted in the urine, and any nickel that is not absorbed from the GI tract is excreted in the feces.

The most prevalent health effect associated with nickel exposure is contact dermatitis in nickel-sensitive individuals. Nickel-related dermatitis results from an allergic reaction to nickel after dermal exposure, and it typically exhibits two components: (1) a simple dermatitis localized in the contact area and (2) chronic eczema or neurodermatitis without apparent connection to such contact. Nickel-related dermatitis has been reported in workers, as well as the general population, after dermal exposure to airborne nickel or liquid nickel solution and also after prolonged contact with metal items containing nickel, such as jewelry and prosthetic devices. Once an individual becomes sensitized to nickel, dermatitis can result from dermal contact with a small amount of nickel or from oral exposure to low doses of nickel. Approximately 10–20% of the general population is sensitized to nickel, with females being more frequently affected than males.

There are very few studies available regarding the oral toxicity of nickel in humans. As noted above, oral exposure to nickel, even at very low doses, can result in flare-ups of

dermatitis in nickel-sensitized individuals. Accidental exposures to much higher doses of nickel in drinking water have resulted in GI upset and neurological symptoms such as headache, giddiness, and weariness.

Noncarcinogenic respiratory effects, such as bronchitis, asthma, and pulmonary fibrosis, have been reported in association with occupational inhalation exposures to nickel, but the evidence for each of these outcomes is limited by a small number of cases and potential coexposures to other sensitizing metals. These effects occurred at nickel concentrations much higher than those found in the environment. A number of mortality studies have indicated that occupational exposure to nickel is not associated with increased mortality from nonmalignant respiratory disease.

Cancers of the lung and nasal sinuses in nickel workers have been described for more than 50 years in association with nickel refining processes (e.g., calcination, smelting, roasting, and electrolysis) and nickel plating and polishing operations (e.g., electrolysis and grinding). These cancers are associated with inhalation exposures to some combination of sulfidic, oxidic, and/or water-soluble nickel or one of these forms of nickel alone. The available data on metallic nickel do not support a role for this form of nickel in carcinogenesis, but it is difficult to determine which particular nickel forms are associated with increased cancer risk because occupational exposures to nickel were almost always to more than one form. Cancers have not been reported in association with background environmental nickel exposure in the general population.

The current oral RfD for nickel (in the form of soluble nickel salts) is 0.02 mg/kg-day, based on decreased body and organ weights in rodents. An inhalation RfD has not been established by US EPA for any form of nickel. The IOM developed a TUL for soluble nickel salts of 1 mg/day for adults, based on decreased body weight gain in rodents.

Occupational exposure guidelines for airborne nickel vary depending on the specific type of nickel compound. The current ACGIH TLV (as an 8-h TWA) for metallic nickel is 1.5 mg/m^3, based on the effects of dermatitis and pneumoconiosis. The current TLV for soluble nickel compounds of 0.1 mg/m^3 is based on lung damage and nasal cancer incidence. The current TLVs of 0.2 mg/m^3 for insoluble nickel compounds and 0.1 mg/m^3 for nickel subsulfide are based on lung cancer incidence. The current TLV for nickel carbonyl is 0.05 ppm and is based on lung and nasal cancer incidence. The current NIOSH-recommended exposure guideline (REL; as a 10-h TWA) for metallic nickel and other nickel compounds is 0.015 mg/m^3 and is based on dermatitis, allergic asthma, and pneumonitis. The current REL for nickel carbonyl is 0.001 ppm, based on CNS effects, pneumonitis, and leukocytosis. The current OSHA PEL (as an 8-h TWA) for metallic nickel and insoluble nickel compounds is 1 mg/m^3, based on dermatitis, asthma, and lung damage. The current PEL for soluble nickel compounds is also 1 mg/m^3, based on dermatitis, asthma, pulmonary effects, eye/nose/throat/skin irritation, and carcinogenic effects. The current PEL for nickel carbonyl is 0.007 mg/m^3, based on lung edema, CNS effects, and suspected teratogenicity.

US EPA has classified nickel refinery dust and nickel subsulfide as class A carcinogens and nickel carbonyl as a class B2 probable human carcinogen for inhalation exposures. The US NTP has classified nickel compounds as known human carcinogens and metallic nickel as reasonably anticipated to be a human carcinogen. IARC has classified nickel compounds as "Group 1: Carcinogenic to humans" and metallic nickel and alloys as "Group 2B: Possibly carcinogenic to humans." Nickel is not presently considered to be a carcinogen by the oral route.

Selenium

Selenium is a naturally occurring element that can be found at background levels in food, soil, and water. It is found primarily in four valence states ($^{2-}$, ^{0}, $^{4+}$, and $^{6+}$). Inorganic selenites ($^{4+}$), such as selenium dioxide and sodium selenite, and selenates ($^{6+}$), such as selenic acid and sodium selenate, are the forms most commonly found in soil and water. Due to their solubility, they are readily taken up by plants and converted to organic selenides, including selenomethionine and selenocysteine.

The maximum selenium concentration in uncontaminated U.S. soils is reported by the United States Geological Survey to be 4.3 mg/kg, while average background concentrations of total selenium in U.S. soils range from 0.25 to 0.53 mg/kg. Selenium concentration in groundwater and surface waters in the United States are generally low and range from 0.0001 to 0.0004 mg/l. Background ambient air concentrations of selenium in the United States range from 0.1 to 10 ng/m^3. Although exposure to selenium may occur through soil, water, and air, the primary route of exposure to selenium in the United States is through food consumption. The average intake of selenium per person in the United States ranges from 0.071 to 0.152 mg/day.

Both inorganic and organic forms of selenium are well-absorbed by the GI tract and are metabolized through a multistep process. When ingested, inorganic selenium is metabolized to the intermediate hydrogen selenide. The selenide may be processed and incorporated into selenoproteins or into transfer ribonucleic acid (tRNA) encoding selenocysteine. Alternatively, the selenide may be methylated and excreted into urine. When ingested, organic forms of selenium found in plants or animal products are sometimes nonspecifically incorporated into tissues such as skeletal muscle, liver, pancreas, stomach, GI mucosa, and erythrocytes. Otherwise, organic forms of selenium can have the same metabolic fate as inorganic selenium. The routes of elimination for selenium are urine, feces, and exhaled breath. Selenium excretion appears to be dose-dependent and related to existing reserves of selenium within the body. In fact,

excretion appears to be a key function in selenium homeostasis, as higher selenium intakes result in a higher percentage of selenium excretion.

Selenium is an essential nutrient for animals (including humans), but excess selenium exposure can induce toxicity (Table 16.2). Selenium plays a role in the development or function of 30 proteins that are required for optimal endocrine and immune responses. Selenium is a cofactor of the glutathione peroxidase family of antioxidant enzymes, which are important molecules in the protection against oxidative stress. Although relationships are not well established, there is some epidemiological evidence that selenium intake is associated with a reduced risk of cancer and cardiovascular disease. Selenium intake is also associated with enhanced immune function. Selenium also acts antagonistically with other metals such as arsenic and mercury, which can ultimately limit the toxicity of those metals.

Despite its potential beneficial attributes, excess selenium can cause adverse effects above a certain dose (~900 µg/day). Acute oral exposure to selenium in humans has also been reported to cause respiratory effects such as pulmonary edema and lung lesions; cardiovascular effects such as tachycardia; GI effects including nausea, vomiting, diarrhea, and abdominal pain; liver effects; and neurological effects such as aches, irritability, chills, and tremors. The potential chronic adverse effects of selenium have been studied based from both environmental exposures and controlled laboratory experiments in humans and animals. In general, chronic oral exposure to selenium in humans is principally associated with dermal and neurological effects. The dermal effects include hair loss, deformation and loss of nails, and discoloration and excessive decay of teeth, while neurological effects include numbness and varying degrees of paralysis. In a series of studies conducted in China, where a population was exposed to high levels of selenium in soil and water due to mining activities, the researchers noted many of the health effects characteristic of selenium toxicity (included a high prevalence of nail deformities, alopecia, skin lesions, tooth decay, and neurological changes (e.g., paresthesias, hyperreflexia)). Based on these studies in China, US EPA developed an RfD for selenium of 5×10^{-3}.

Based on both human and animal investigations, the evidence supporting a relationship between environmental selenium and cancer is limited. Neither US EPA nor IARC has classified selenium as a carcinogen. Although most forms of selenium that are commonly encountered by humans are not carcinogenic, in animal studies (rodents, specifically), selenium sulfide and ethyl selenac (selenium diethyldithiocarbamate) administered orally induce tumors. Based on this result, US EPA has classified selenium sulfide as a Group B2 probable human carcinogen.

Interestingly, in populations experiencing selenium deficiency, an increased cancer risk has been observed. The relationship between selenium supplementation and cancer prevention is more complicated and under active investigation. Overall, a cancer health benefit associated with selenium supplementation has been noted in individuals with a low selenium status, but supplementation in selenium-sufficient individuals offers limited benefit. Selenium has also been found to have chemopreventive qualities for people with existing tumors.

In addition to the RfD established by US EPA, an upper tolerable level of 400 µg/day has been established for selenium by IOM; an intake of 5 µg/day is recommended for adults. An RfC has not been established. US EPA has established an MCL of 0.05 mg/l (50 µg/l) for selenium in U.S. drinking water supplies. Occupational exposure guidelines have also been developed for various selenium compounds. ACGIH's 8-h TWA, NIOSH's TWA, and OSHA's PEL for selenium is 0.2 mg/m^3.[2]

Thallium

Thallium naturally occurs in trace amounts throughout the environment in two valence states, each with distinct chemical properties with regard to solubility and bioavailability. Thallium compounds in the monovalent state are more stable (and common) in the environment than the trivalent forms. The natural content of thallium in soil ranges from 0.1 to 3 mg/kg, with an average of about 1 mg/kg or less. In a U.S. survey of thallium concentration in drinking water sources, thallium was detected in only 0.68% of the samples with an average concentration of 0.89 µg/l. Information on the amount of thallium in air is limited, but in six U.S. cities, US EPA found that levels typically ranged from 0.02 to 0.1 ng/m^3.

Thallium is readily absorbed into the body through multiple exposure routes. Once absorbed, thallium is rapidly distributed to all organs with an apparent blood half-life of less than 5 min. Animal evidence suggests that thallium crosses the blood–brain and placental barriers to some extent. The primary routes of thallium excretion for animals and humans are the urine and feces, although some excretion may also occur through hair and sweat.

Several health effects have been observed following acute oral thallium exposure, mainly from incidents involving intentional or accidental poisonings. Also, thallium was historically used medicinally at high concentrations to treat skin conditions and the side effects of tuberculosis. In general, higher doses are associated with immediate death secondary to cardiac or respiratory failure. Sublethal acute doses of thallium are associated with GI problems, such as abdominal pain, vomiting, diarrhea, and nausea. Neurological deficits can also develop, including paresthesias, limb weakness, and abnormal mental function. One of the most characteristic signs of both acute and chronic thallium toxicity is hair loss.

[2] Note that alternative exposure guideline values exist for selenium hexafluoride and hydrogen selenide.

Symptoms of chronic oral thallium toxicity are similar to acute effects, although symptoms are delayed. Information on chronic oral thallium toxicity is mainly from a few isolated studies on populations living in areas of localized thallium releases from nearby industries, where exposure comes from the uptake of thallium into crops. The most common symptoms in these exposed communities were alopecia, peripheral neuropathy, visual deficits, and GI problems. Treatment with a chelating agent led to a full recovery in most individuals. Two of the more effective chelating agents for thallium are activated charcoal and Prussian Blue.

Limited data are available on chronic thallium exposure via inhalation. The only available studies are occupational and usually involved exposure to multiple chemicals. Thus, isolating the effect from thallium, specifically, is often problematic. Nonetheless, workers in the glass production and cement industries, with presumed elevated exposures to thallium, reported symptoms consistent with oral thallium exposure including GI effects, leg pain, fatigue, alopecia, and psychological deficits. Thallium is not recognized to cause adverse effects via the dermal route.

No human or animal studies demonstrate that thallium causes cancer, and there is no convincing evidence that thallium is mutagenic. Some genotoxicity reports exist for thallium, but many of the experiments involved coexposure to other compounds. Based on the lack of evidence for carcinogenicity, US EPA has determined that thallium compounds are "not classifiable as to human carcinogenicity."

Based on the available information on thallium's noncancer toxicology, US EPA developed oral RfDs for several compounds containing thallium ranging from 7×10^{-5} to 9×10^{-5}. These values, however, were withdrawn in 2009 when US EPA concluded that the underlying toxicity data had too many limitations to derive a reliable value. An RfC is not published on IRIS. US EPA has established an MCLG of 0.0005 mg/l (0.5 µg/l) and an MCL of 0.002 mg/l (2 µg/l) for thallium. For occupational exposures to soluble thallium compounds (as Tl), OSHA, ACGIH, and NIOSH have set a TWA exposure level of 0.1 mg/m^3, while NIOSH has also developed an immediately dangerous to life and health criterion of 15 mg/m^3.

Zinc

Zinc is found in natural form in the air, soil, and water and occurs in the environment primarily in the Zn^{2+} oxidation state. It is a bluish-white metal that can combine with a number of other elements. The most common zinc ore is zinc sulfide. In water, zinc partitions to sediments or suspended solids through sorption. Zinc is likely to sorb strongly onto soil, and the mobility of zinc in soil depends on the solubility of the speciated forms of the compound and on soil properties such as pH, redox potential, and cation exchange potential. In air, zinc is present primarily as small dust particles. Humans are exposed to small amounts of zinc in food and drinking water each day. Levels in air are generally low and fairly constant. Occupational exposure to zinc occurs in a number of mining and industrial activities, such as the manufacture of zinc-containing alloys, paints, and pesticides.

Approximately 20–30% of ingested zinc is absorbed due to homeostatic regulation (i.e., passive and carrier-mediated diffusion processes). Zinc is distributed throughout the body, and high concentrations are found in muscle tissues, bone, prostate, pancreas, liver, and kidneys. Zinc is also absorbed following inhalation exposure, and retention is dependent on the form and particle size. Ingested zinc (70–80%) is excreted in both the feces and the urine. Zinc induces metallothionein synthesis, and absorption decreases when metallothionein is saturated.

Zinc is an essential trace element in the diet, plays a role in numerous enzymatic reactions, and is critical for normal growth and development (Table 16.2). Zinc deficiency can have numerous adverse effects on the normal function of all of these systems. The RDA for zinc ranges from 2–3 mg/day for infants to 8–12 mg/day for adolescents and adults to 11–13 mg/day for pregnant or lactating women.

Zinc salts may be irritating to the GI tract, and when ingested, they may act as emetics. Excessive zinc exposure in humans is uncommon; however, it may result in abdominal cramps, vomiting, nausea, gastric irritation, anemia, or decreased cholesterol levels. Zinc exposure also plays a role in copper absorption and may result in copper deficiency. Zinc does not appear to increase cancer incidence, and there is actually some evidence that it decreases the cancer effects from other toxic metals (e.g., cadmium).

Inhalation of zinc fumes is commonly associated with metal fume fever and symptoms of fever, chills, shakes, headache, weakness, and a metallic taste in the mouth. Symptoms typically occur within hours of exposure and may last for 1–2 days, and a tolerance may develop after repeated exposure. Metal fume fever is thought to be an immune response characterized by flu-like symptoms and impaired lung function. Similar health effects are found in animal studies. Development of cancer has not been associated with inhalation exposure to zinc.

Zinc chloride salts are corrosive to the skin at levels less than 0.5 mg/cm^2, while zinc sulfate is less irritating and zinc oxide does not appear to cause irritation. Since the skin is relatively insensitive to zinc oxide, it is a common ingredient in sunscreens. Ocular effects from exposure to zinc may cause conjunctivitis, lacrimation, or in extreme cases, corneal damage. Most symptoms may abate after several weeks, while residual effects may persist for a long period of time.

US EPA has established a daily oral RfD of 3×10^{-1} mg/kg for zinc; however, no inhalation RfC has been established. Under the 2005 US EPA Guidelines for Carcinogenic Risk Assessment, data for zinc are inadequate to evaluate human

carcinogenicity. Occupational exposure values have been developed by OSHA for zinc chloride fume (TWA-PEL = 1 mg/m^3) and zinc oxide (TWA-PEL = 5 mg/m^3 for respirable fraction). ACGIH has set TWA-TLVs of 1 mg/m^3 for zinc chloride fume and 2 mg/m^3 (respirable fraction) for zinc oxide. US EPA has not set a primary drinking water MCL for zinc; however, a secondary standard (based on odor and taste) was established at 5 mg/l.

Rare Earth Elements (Lanthanides)

REEs (lanthanides) are a group of 15 metals that include cerium (Ce), dysprosium (Dy), erbium (Er), europium (Eu), gadolinium (Gd), holmium (Ho), lanthanum (La), lutetium (Lu), neodymium (Nd), praseodymium (Pr), promethium (Pm), samarium (Sm), terbium (Tb), thulium (Tm), and ytterbium (Yb). The elements scandium (Sc) and yttrium (Y) are often included in this group. The common atomic feature of most lanthanides is that they fill the 4*f* electronic subshell, which limits variation in chemical properties. REEs are considered rare due to complex extraction processes from mineral ores. Except for promethium (which only has radioactive isotopes), REEs are generally more abundant in the earth's crust relative to other toxic metals (e.g., cadmium and lead). For example, the crustal abundance ranges from less than 1 to 1 ppm (e.g., europium, holmium, lutetium, terbium, and thulium) to 5–70 ppm (e.g., Ce, La, Nd, and Y).

Lanthanides typically form ions in the REE^{3+} oxidation state, while some are stable in the $^{2+}$ (Eu^{2+} and Yb^{2+}) or $^{4+}$ (Ce^{4+} and Tb^{4+}) states. These metals react with oxygen to form oxides with the formula REE_2O_3 (except for cerium, which forms CeO_2). All REEs have relatively high melting points (>700 to >1500 °C). They are generally soluble in water as sulfates, nitrates, and chlorates and insoluble as phosphates, fluorides, and oxalates. REEs are often grouped into two broad classes due to their co-occurrence in mineral ores and chemical similarities. "Light" REEs include lanthanum, cerium, praseodymium, neodymium, samarium, europium, and gadolinium, while "heavy" REEs include yttrium, terbium, dysprosium, holmium, erbium, thulium, ytterbium, and lutetium. The utility of these elements in technology and electronics is rapidly expanding. Common uses include formation of metal alloys, lighting and chemical catalysts, glass and ceramics, phosphors and electronics, nanomaterials, pharmaceuticals, and medical applications. REEs are also critical elements needed for clean energy technology applications and are used in fuel additives to reduce carbon emissions and in magnets and batteries for wind turbines and hybrid or electrical vehicles. As the demand for REEs grows, and the level of production increases, the potential for environmental releases and exposures similarly rises.

Absorption of lanthanides is highly dependent on solubility. Radioisotope distribution studies indicate that REEs are minimally absorbed orally (1% of administered dose) while almost completely absorbed following injection. The primary deposition sites include the skeleton, liver, and kidney from oral or intramuscular exposure, while inhalation studies suggest REEs are deposited in the lung tissues and over time translocated to the skeleton, liver, or other organs. Light REEs are predominantly excreted in the feces, while heavy REEs are excreted in the urine.

Few epidemiological studies have examined the association between oral exposure to REEs and health effects. In areas with elevated environmental concentrations of REEs (particularly cerium), some epidemiological studies have suggested an increased incidence of arteriosclerosis, endomyocardial fibrosis, or acute myocardial infarction. Animal studies have not been conducted to further investigate these findings. Acute oral toxicity of lanthanides is generally low due to poor GI absorption (LD_{50} > 1300 to >5000 mg/kg). Subchronic and chronic studies of individual REEs or mixtures have demonstrated that the liver and kidney are primary target organs. Animal models have generally provided negative results with respect to reproductive, developmental, or carcinogenic effects. Recent investigations have focused on the potential neurotoxicity of REEs. In an epidemiological report, children living in a rare earth mining district in China had higher body burdens of REEs and significantly lower IQ scores. Rodent models have also demonstrated reduced performance on learning and memory tests following oral exposure to REEs. However, the available toxicological data are insufficient to fully characterize the dose–response relationship.

Chronic exposure to lanthanide dusts in occupational settings has been associated with the development of pneumoconiosis (i.e., inflammation, cough, and fibrosis from exposure to metallic dusts). This is based primarily on clinical cases of workers exposed to cerium oxide and rare earth oxides in fumes generated from carbon-arc lamps. In these cases, accumulation of cerium particles leading to pulmonary overload has been suggested to account for the observed effects; however, further information is needed to validate this proposed mode of action. Few long-term exposure studies have been conducted using animal models. A 13-week study examined the effects from inhalation of ceric oxide in rats and an lowest observed adverse effect level (LOAEL) of 5 mg/m^3 was reported based on an increased incidence of lymphoid hyperplasia in the bronchial lymph nodes. There is inadequate information to evaluate the carcinogenicity of lanthanides via inhalation. Due to the increasing industrial uses of REEs in technology and industry, particularly for nanomaterials, additional studies are needed to evaluate exposure and effect from the inhalation route.

The toxic action of lanthanides on the skin has not been well described. Absorption of REEs by healthy skin is negligible and by abraded skin is limited. Mild to severe skin irritation was observed for a subset of REEs in rabbits and guinea pigs using the Draize procedure. Conjunctivitis

occurred in rabbits following direct ocular application of REE chlorides (dysprosium, erbium, europium, and holmium).

Due to the emerging application of REEs and the relative lack of toxicological information, regulatory standards have not been prepared for a majority of these metals. US EPA has developed an RfC for cerium oxide and compounds of 9×10^{-4} mg/m^3 based on increased incidence of alveolar epithelial hyperplasia in the lungs of rats. Inadequate evidence is available to evaluate the carcinogenicity of REEs. Occupational exposure guidelines are not available for REEs, except for yttrium (OSHA TWA-PEL and ACGIH TWA-TLV are 1 mg/m^3). Drinking water standards have not been established for these elements.

16.8 SUMMARY

This chapter has presented an overview of relevant physical/chemical and toxicological considerations for metals, relevant to understanding the health effects of metals. Perhaps, the most important conclusion from this discussion of "general" properties is that one cannot generalize across metals as a group or even among different forms (e.g., organic vs. inorganic) or valence states of the same metal. Half-lives in the body, targets of accumulation, and excretion routes vary dramatically across metals and among different species of the same metal as do the nature of effects, from the molecular to the organismic level. Some metals, for example, copper and zinc, are essential nutrients, adding additional complexity to evaluating the toxicity of these specific metals. In addition, variations in host susceptibility, not only across population groups (e.g., lead and child vs. adult susceptibility) but also within populations (e.g., copper and individuals with or without Wilson's disease), are also important factors that influence toxicity of metals.

The following bibliography provides additional source material on the toxicity of metals.

ACKNOWLEDGMENTS

The authors would like to thank the following scientists from Gradient for their contributions to this manuscript: David G. Dodge, Tom A. Lewandowski, Robyn L. Prueitt, Karrie A. Radloff, and Mara R. Seeley. We would also like to thank Steven G. Donkin, Danny L. Ohlson, and Christopher M. Teaf who authored the previous edition of this chapter. Their work was used as a basis for this updated chapter.

REFERENCES AND SUGGESTED READING

Agency for Toxic Substances and Disease Registry (ATSDR). ToxProfiles 2010: Including ToxFAQs and Public Health Statements (2 CDs). Agency for Toxic Substances and Disease Registry, Atlanta; 2010.

Andersen ME, Dorman DC, Clewell HJ III, Taylor MD, Nong A. Multi-dose-route, multi-species pharmacokinetic models for manganese and their use in risk assessment. J Toxicol Environ Health A 2010;73:217–234.

Aschner M, Ceccatelli S. Are neuropathological conditions relevant to ethylmercury exposure? Neurotox Res 2010;18:59–68.

Aschner M, Syversen T. Methylmercury: recent advances in the understanding of its neurotoxicity. Ther Drug Monit 2005;27:278–283.

Bakir F, Damluji SF, Amin-Zaki L, Murtadha M, Khalidi A, al-Rawi NY, Tikriti S, Dahahir HI, Clarkson TW, Smith JC, Doherty RA. Methylmercury poisoning in Iraq. Science 1973;181:230–241.

Barceloux DG. Molybdenum. J Toxicol Clin Toxicol 1999;37:231–237.

Bates MN. Mercury amalgam dental fillings: an epidemiologic assessment. Int J Hyg Environ Health 2006;209:309–316.

Bersenyi A, Berta E, Kadar I, Glavits R, Szilagyi M, Fekete SG. Effects of high dietary molybdenum in rabbits. Acta Vet Hung 2008;56:41–55.

Caldwell KL, Mortensen ME, Jones RL, Caudhill SP, Osterloh JD. Total blood mercury concentrations in the U.S. population: 1999–2006. Int J Hyg Environ Health 2009;212:588–598.

Castoldi AF, Johansson C, Onishchenko N, Coccini T, Roda E, Vahter M, Ceccatelli S, Manzo L. Human developmental neurotoxicity of methylmercury: impact of variables and risk modifiers. Regul Toxicol Pharmacol 2008;51:201–214.

Clarkson TW. Environmental contaminants in the food chain. Am J Clin Nutr 1995;61:682S–686S.

Clarkson TW, Magos L. The toxicology of mercury and its chemical compounds. Crit Rev Toxicol 2006;36:609–662.

Ekino S, Susa M, Ninomiya T, Imamura K, Kitamura T. Minamata disease revisited: an update on the acute and chronic manifestations of methyl mercury poisoning. J Neurol Sci 2007;262:131–144.

Flora SJ, Pachauri V. Chelation in metal intoxication. Int J Environ Res Public Health 2010;7:2745–2788.

Gohfeld M. Cases of mercury exposure, bioavailability, and absorption. Ecotoxicol Environ Saf 2003;56:174–179.

Goodman JE, Prueitt RL, Thakali S, Oller AR. The nickel ion bioavailability model of the carcinogenic potential of nickel-containing substances in the lung. Crit Rev Toxicol 2011;41:142–174.

Haber LT, Erdreicht L, Diamond GL, Maier AM, Ratney R, Zhao Q, Dourson ML. Hazard identification and dose response of inhaled nickel-soluble salts. Regul Toxicol Pharmacol 2000;31:210–230.

He X, Zhang Z, Zhang H, Zhao Y, Chai Z. Neurotoxicological evaluation of long-term lanthanum chloride exposure in rats. Toxicol Sci 2008;103:354–361.

Hirano S, Suzuki KT. Exposure, metabolism, and toxicity of rare earths and related compounds. Environ Health Perspect 1996;104:85–95.

Institute of Medicine (IOM). Dietary Reference Intakes for Vitamin C, Vitamin E, Selenium, and Carotenoids. *Standing Committee on the Scientific Evaluation of Dietary Reference*

Intakes, Food and Nutrition Board. National Academy Press, Washington; 2000.

Institute of Medicine (IOM). Seafood Choices, Balancing Benefits and Risks. *Committee on Nutrient Relationships in Seafood: Selections to Balance Benefits and Risks*. National Academies Press, Washington; 2006.

International Committee on Nickel Carcinogenesis in Man. Report of the International Committee on Nickel Carcinogenesis in Man. Scand J Work Environ Health 1990;16:1–82.

Johansson C, Castoldi AF, Onishchenko N, Manzo L, Vahter M, Ceccatelli S. Neurobehavioural and molecular changes induced by methylmercury exposure during development. Neurotox Res 2007;11:241–260.

Jones RL, Homa DM, Meyer PA, Brody DJ, Caldwell KL, Pirkle JL, Brown MJ. Trends in blood lead levels and blood lead testing among US children aged 1 to 5 years, 1988–2004. Pediatrics 2009;123:e376–e385.

Langford NJ, Ferner RE. Toxicity of mercury. J Hum Hypertens 1999;13:651–656.

Lewandowski T. Plausibility of toxicological interaction between lead and methylmercury. Toxicol Environ Chem 2011;93: 1423–1462.

Mendel RR. Molybdenum: biological activity and metabolism. Dalton Trans 2005;21:3404–3409.

Menezes-Filho JA, Bouchard M, Sarcinelli PDN, Moreira JC. Manganese exposure and the neuropsychological effect on children and adolescents: a review. Rev Panam Salud Publica 2009;26:541–548.

National Research Council (NRC). *Toxicological Effects of Methylmercury*. Washington: National Academy Press; 2000.

Novotny JA, Turnlund JR. Molybdenum intake influences molybdenum kinetics in men. J Nutr 2007;137:37–42.

Parker S, Todd J, Schwartz B, Pickering L. Thimerosal-containing vaccines and autistic spectrum disorder: a critical review of published original data. Pediatrics 2005;115:200.

Peter ALJ, Viraraghavan T. Thallium: a review of public health and environmental concerns. Environ Int 2005;31:493–501.

Petito Boyce C, Lewis AS, Sax SN, Eldan M, Cohen SM, Beck BD. Probabilistic analysis of human health risks associated with background concentrations of inorganic arsenic: use of a margin of exposure approach. Hum Ecol Risk Assess 2008;14:1159–1201.

Roman HA, Walsh TL, Coull BA, Dewailly E, Guallar E, Hattis D, Mariën K, Schwartz J, Stern AH, Virtanen JK, Rice G. Evaluation of the cardiovascular effects of methylmercury exposures: current evidence supports development of a dose-response function for regulatory benefits analysis. Environ Health Perspect 2011;119:607–614.

Santamaria AB. Manganese exposure, essentiality and toxicity. Indian J Med Res 2008;128:484–500.

Schroeter JD, Nong A, Yoon M, Taylor MD, Dorman DC, Andersen ME, Clewell HJ III. Analysis of manganese tracer kinetics and target tissue dosimetry in monkeys and humans with multi-route physiologically based pharmacokinetic models. Toxicol Sci 2011;120:481–498.

Schultz ST. Does thimerosal or other mercury exposure increase the risk for autism? A review of current literature. Acta Neurobiol Exp (Wars) 2010;70:187–195.

Stern AH. A review of the studies of the cardiovascular health effects of methylmercury with consideration of their suitability for risk assessment. Environ Res 2005;98:133–142.

U.S. Geological Survey. Major- and Trace-Element Concentrations in Soils from Two Continental-Scale Transects of the United States and Canada. *USGS Open-File Report 2005-1253*. US Geological Survey, Reston; 2005

U.S. Geological Survey. Quality of Water from Domestic Wells in Principal Aquifers of the United States, 1991–2004. *USGS Scientific Investigations Report 2008-5227*. U.S. Geological Survey, Reston; 2009.

Vyskocil A, Viau C. Assessment of molybdenum toxicity in humans. J Appl Toxicol 1999;l19:185–192.

Yoon M, Schroeter JD, Nong A, Taylor MD, Dorman DC, Andersen ME, Clewell HJ III. Physiologically based pharmacokinetic modeling of fetal and neonatal manganese exposure in humans: describing manganese homeostasis during development. Toxicol Sci 2011;122:297–316.

17

PROPERTIES AND EFFECTS OF PESTICIDES

JANICE BRITT

The term "pesticide" encompasses a group of chemical compounds that are used for the elimination or control of pests. Pesticides are grouped into classes based on their target of action. They comprise a diverse group as can be seen in Table 17.1.

Pesticides have both economic and public health benefits. They are used for the control of vector-borne diseases (e.g., malaria, Lyme disease, West Nile Virus, avian flu, and Rocky Mountain spotted fever), for the promotion of agricultural production in the United States as well as in other countries, and by homeowners for the control of domestic pests (e.g., household and garden pests). More recently, agents such as vaporized hydrogen peroxide, methyl bromide, paraformaldehyde, and ethylene oxide have been approved as pesticides to be used against biological agents such as *Bacillus anthracis* (anthrax-producing bacteria that caused threats to national security in the United States).

Individuals may be exposed to pesticides either occupationally (e.g., from working in a pesticide formulating plant or from commercial pesticide application) or environmentally (e.g., from food products such as fruits and vegetables treated for pests). Individuals may also be exposed to pesticides at their residences (e.g., from use as home or garden insecticide). The U.S. Environmental Protection Agency (USEPA) has recently reported the most commonly used pesticides (see Table 17.2).

The registration and regulatory requirements concerning pesticides are governed under the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA). Rules aimed at the protection of agricultural and greenhouse workers who used pesticides were passed in 1992 by the USEPA and are located in 40 Code of Federal Regulations Parts 156 and 170. Occupational Safety and Health Administration (OSHA) as well as the American Conference of Governmental Industrial Hygienists (ACGIH) also publish guidelines for occupational exposures to pesticides in air. For some pesticides, the USEPA Integrated Risk Information System (IRIS) and/or the Office of Pesticide Programs (OPP) has set various guidelines for pesticides for drinking water (e.g., maximum contaminant levels and Human Health Benchmarks for Pesticides (HHBPs)) or in air (e.g., reference concentration (RfC)) as well as estimates of cancer risks (e.g., oral and/or inhalation slope factors).

This chapter will discuss the classes of the most commonly used pesticides and will include a discussion of the following with respect to these pesticide classes:

- Uses
- Mechanism of action
- Pharmacokinetics
- Acute and chronic effects from exposure
- Biological monitoring
- Treatment of pesticide overexposure
- Regulatory information

17.1 ORGANOPHOSPHATE AND CARBAMATE INSECTICIDES

Introduction, Use, and History of Organophosphates and Carbamates

In the past, organophosphate compounds were used as replacements for the more persistent organochlorine insecticides because they did not bioaccumulate in tissues and

Principles of Toxicology: Environmental and Industrial Applications, Third Edition. Edited by Stephen M. Roberts, Robert C. James and Phillip L. Williams.

TABLE 17.1 Classes of Pesticides and Their Targets

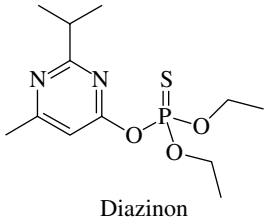
Class	Target
Acaricides	Mites and ticks
Algicide	Algae
Antifouling agents	Organisms that attach to underwater surfaces
Antimicrobials	Bacteria and viruses
Attractant	Attracts pests
Avicides	Birds
Disinfectants/ sanitizers	Kills or inactivates microorganisms on inanimate objects
Fumigants	Pests in buildings or soil
Fungicides	Fungi
Herbicides	Weeds or undesirable plants
Insecticides	Insects
Larvicides	Larvae such as mosquitoes
Miticides	Mites
Molluscicides	Slugs and snails
Nematicides	Nematodes
Ovicides	Eggs of pests (e.g., fleas)
Piscicide	Fish
Repellants	Repels pests such as mosquitoes
Rodenticides	Rats, mice, and other rodents
Silvicide	Trees, shrubs, or forests
Termiticide	Termites and ants

Source: Adapted from Delaplane (1996).

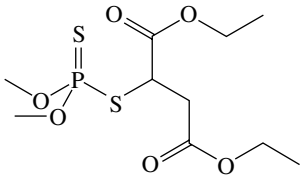
TABLE 17.2 Most Commonly Used Pesticides in the United States

Most Commonly Used Pesticides in U.S. Agricultural Market Sector
Glyphosate
Atrazine
Metam sodium
Metolachlor-S
Acetochlor
Most Commonly Used Pesticides in Non-Agricultural Sectors of the United States
Home and Garden Market
2,4-Dichlorophenoxyacetic acid
Glyphosate
Carbaryl
MCPP
Pendimethalin
Industrial/commercial/government uses
2,4-Dichlorophenoxyacetic acid
Glyphosate
Chlorothalonil
MSMA
Diuron

organisms or accumulate in the environment. For example, chlorpyrifos, an organophosphate compound, was at one time, a widely used termiticide, serving as a substitute for the more persistent organochlorine compounds that were used in the past. However, several years ago, due to concern over toxicity of chlorpyrifos, the USEPA curtailed the use of this compound, with a few exceptions. Examples of organophosphate insecticides (and their active ingredient) include Dursban™ (chlorpyrifos), Knox Out™ (diazinon), and Vapona™ (dichlorvos) (see Figure 17.1 for examples of organophosphate insecticides). Examples of carbamate pesticide products (and their active ingredient) commonly used are Sevin™ (carbaryl) and Temik™ (aldicarb) (see Figure 17.2 for examples of carbamate insecticides). It should be noted that organophosphate compounds are not only used as pesticides; chemicals in this class are also used as therapeutic agents for the treatment of glaucoma and myasthenia gravis in humans. For example, the organophosphate echothiophate iodide is used to treat glaucoma. Rivastigmine (Exelon®), a reversible cholinesterase inhibitor, is currently being used to treat dementia. Currently, clinical investigators are exploring the possible use of other cholinesterase inhibitors to treat Alzheimer's disease, Parkinson's disease dementia, dementia with Lewy bodies, and other medical conditions.

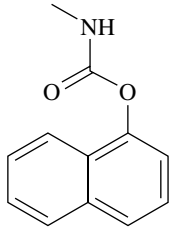


Diazinon

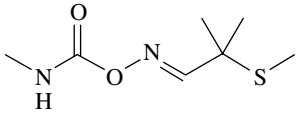


Malathion

FIGURE 17.1 Examples of organophosphate insecticides.



Carbaryl



Aldicarb

FIGURE 17.2 Examples of carbamate insecticides.

Tetraethyl pyrophosphate (TEPP), the first anticholinesterase compound, was synthesized in the mid-1800s by the French chemist Philippe de Clermont and the Russian chemist Wladimir Moshnin who were students in Adolphe Wrutz's laboratory in Paris, France. Schrader in Germany, who was developing toxic OP compounds during WWII, uncovered the insecticidal properties that TEPP possessed. While TEPP was an effective insecticidal compound, it was unstable in the environment, and so there was an effort to develop more stable compounds. This effort resulted in the development of organophosphate insecticides, including dimefox in 1940, schradan in 1942, and parathion in 1944.

Mechanism of Action

Both organophosphate and carbamate classes of compounds have the same mechanism of action in insects as well as in mammals (including humans)—the inhibition of the enzyme acetylcholinesterase. The inhibition of acetylcholinesterase by organophosphates and carbamates compounds is the mechanism of action that is responsible for the acute symptomatology associated with these compounds.

Acetylcholinesterase is an enzyme located in the synaptic cleft, and its function is the breakdown of acetylcholine, the neurotransmitter present at the following sites: postganglionic parasympathetic nerves, somatic motor nerves endings in skeletal muscle, preganglionic fibers in the parasympathetic and sympathetic nerves, and in some synapses in the central nervous system. Organophosphate and carbamate insecticides act by inhibiting the enzyme acetylcholinesterase at its esteratic site, resulting in an accumulation of the neurotransmitter acetylcholine in nerve tissue and at the effector organ. This accumulation then results in the continued stimulation of cholinergic synapses and at sufficient level leads to the signs and symptoms associated with overexposure to these compounds (discussed later in this section).

Absorption and Metabolism

Organophosphates and carbamates can be readily absorbed via ingestion, dermal, and inhalation routes due to their lipophilic nature. Within the class of organophosphate insecticides, there are direct organophosphate inhibitors (those containing =O) and organophosphate indirect inhibitors (those containing =S), depending on whether or not they require metabolic activation before they can inhibit acetylcholinesterase. In other words, the indirect organophosphate compounds (containing =S) must undergo bioactivation to become biologically active (containing =O). The indirect inhibiting compounds, including organophosphates such as parathion, diazinon, malathion, and chlorpyrifos, become more toxic than the parent compound upon metabolism. In the case of these indirect inhibitors, oxidative desulfuration (replacement of the sulfur atom with an oxygen atom as described earlier) results in the formation of the oxon of the parent compound (e.g., parathion → paraoxon, diazinon → diazoxon, malathion → maloxon, and chlorpyrifos → chlorpyrifos-oxon). This metabolism occurs via the mixed function oxidase system of the liver.

Organophosphate insecticides not requiring metabolic activation (i.e., direct inhibitors, discussed later) can produce local toxic effects at the site of exposure such as sweating (dermal exposure), miosis or pinpoint pupils (eye contact), and/or bronchospasms (inhalation exposure). Both the organophosphate and carbamate insecticides have relative short biological half-lives and are fairly rapidly metabolized and excreted.

Once cholinesterase activity has been inhibited in the body by an organophosphate compound, the recovery of that compound is dependent on the reversal of inhibition, aging, and the rate of regeneration of a new enzyme. A chemical reaction that organophosphate insecticides can undergo in the body once they are bound to the cholinesterase enzyme is called "aging." Aging involves the dealkylation of the compound once it is bound to the cholinesterase enzyme. In this "aged" form, the organophosphate compound is tightly bound to the enzyme and will not release itself from the enzyme. Once the aging reaction has occurred, treatment with medications (such as pralidoxime, discussed later) is not effective. Once the cholinesterase molecule has been irreversibly inhibited via the aging process, the only manner in which the enzyme activity may be restored is through synthesis of new enzyme.

In addition to the aging reaction, organophosphates can undergo various phase I and II biotransformation pathways including oxidative, hydrolytic, GSH-mediated transfer and conjugation reactions.

Carbamate compounds do not require metabolic activation in order to inhibit cholinesterase. Carbamate insecticides are not considered irreversible inhibitors like some organophosphate insecticides. Cholinesterase inhibition by carbamate compounds is readily reversible, with reversal of inhibition occurring typically within a few hours after exposure. This rapid reversal of the cholinesterase enzyme activity leads to a much shorter duration of action and thus a shorter period of intoxication than that seen in cases of organophosphate overexposure. Carbamates do not undergo "aging" as do the organophosphates. As with the organophosphates, carbamates also can undergo various phase I and II metabolism reactions.

Acute Effects of Organophosphate and Carbamate Insecticides

The effects of organophosphate and carbamate insecticides can be either local (e.g., sweating from localized dermal exposure as mentioned earlier) or systemic. Signs and symptoms of overexposure to organophosphate and carbamate

compounds occur fairly rapidly after exposure, with effects typically seen beginning from 5 min to 12 h after exposure. A diagnosis of organophosphate intoxication typically is based on an exposure history of 6 h or less before the onset of signs and symptoms. It has been suggested that if symptoms appear more than 12 h after the exposure, then another etiology should be considered, and if the symptoms begin 24 h after the exposure, then organophosphate intoxication should be considered to be equivocal.

Symptoms of carbamate overexposure generally develop within 15 min to 2 h of exposure and typically last only several hours, a duration much shorter than that of the typical overexposure to organophosphate pesticide. Symptoms that are present 24 h following exposure are likely not a result of overexposure to carbamate insecticides.

The acute signs and symptoms seen in cases of overexposure to both organophosphate and carbamate pesticides are related to the degree of inhibition of acetylcholinesterase in the individual. The clinical manifestation of overexposure to organophosphate and carbamate compounds is a result of muscarinic, nicotinic and central nervous system symptoms (see Table 17.3). In systemic intoxications with organophosphate and carbamate compounds, the muscarinic effects are generally the first effects to develop.

Organophosphate intoxication is diagnosed on clinical suspicion; an opportunity for exposure to the compound; odor of pesticides; signs and symptoms consistent with organophosphate overexposure; and significant inhibition (i.e., 50% inhibition) of cholinesterase enzyme as measured in the plasma and/or in red blood cells (this will be discussed later). Signs and symptoms resulting from overexposure to this organophosphate and carbamate compounds can be best described by the mnemonic DUMBELS: *D*iarrhea, *U*rination, *M*iosis (pinpoint pupils), *B*ronchospasm, *E*mesis (vomiting), *L*acrimation (tearing), and *S*alivation. Signs and symptoms associated with exposure to organophosphates and carbamate compounds generally do not occur unless acetylcholinesterase activity is approximately 50% or less of normal activity.

Signs and symptoms in cases of mild to moderate organophosphate intoxication typically resolve within days to weeks following exposure. In cases of severe organophosphate intoxication, it can be 3 months or so before cholinesterase red blood cell levels return to normal. Death from organophosphate intoxication is usually due to respiratory failure from depression of the respiratory center in the brain, paralysis of the respiratory muscles, and excessive bronchial secretions, pulmonary edema, and bronchoconstriction. Death in individuals with acute organophosphate intoxication that are untreated typically occur within the first 24 h, and within 10 days in treated individuals. If there is no anoxia, complete recovery will occur, in general, within about 10 days after the exposure incident.

Carbamate intoxication presents similar to that of organophosphate intoxication. Cases of carbamate intoxication resolve much more quickly than cases of organophosphate overexposure due to the rapid reversal of acetylcholinesterase enzyme with the carbamate insecticides as well as to the rapid biotransformation *in vivo*.

Chronic Effects of Organophosphate and Carbamate Insecticides

In general, the main reported chronic effect that may result from exposure to organophosphate insecticides is delayed neuropathy. Organophosphate-induced delayed neuropathy (OPIDN) has been associated with exposure to only a few

TABLE 17.3 Signs and Symptoms of OP Intoxication

Muscarinic (Parasympathetic Nervous System)	Nicotinic (Sympathetic and Motor Nervous Systems)	Nicotinic and Muscarinic (Central Nervous System)
Increased salivation	Sweating	Giddiness
Increase lacrimation	Pallor	Tension
Bronchospasm	Hypertension	Anxiety
Restless	Muscle fasciculations	Restlessness
Dyspnea	Muscle cramps	Insomnia
Anorexia	Motor weakness	Nightmares
Nausea	Tachycardia (↑ heart rate)	Headache
Vomiting	Paralysis	Tremors
Abdominal cramps		Drowsiness
Diarrhea		Confusions
Miosis (pinpoint pupils)		Slurred speech
Blurred vision		Ataxia
Urinary frequency and incontinence		Cheyne–Stokes respiration
Bradycardia (↓ heart rate)		Convulsions
Hypotension		Coma
Cyanosis		

organophosphate compounds, with cases of OPIDN having occurred almost exclusively at near-lethal exposure levels. No permanent effects generally result from carbamate intoxication; delayed neuropathy does not occur as a result of carbamate poisoning (see discussion later).

Organophosphate-Induced Delayed Neuropathy

A few of the organophosphates have been associated with the development of a rare condition—a delayed predominantly motor peripheral neuropathy, termed OPIDN. In the United States in the 1930s, individuals developed OPIDN, also called "Ginger Jake" paralysis, after consuming ginger liquor contaminated with triorthyl cresyl phosphate (TOCP). Other outbreaks of OPIDN have occurred in relation to the consumption of cooking oil contaminated with TOCP. Organophosphates that have been associated with OPIDN include TOCP, mipafox, trichlorphon, leptophos, and methamidophos. It should be pointed out that only a few of the organophosphate compounds actually are capable of causing OPIDN.

The development of OPIDN is not physiologically related to cholinesterase inhibition. The nerve lesion in OPIDN is that of a distal symmetric predominantly motor polyneuropathy of the long, large diameter axons (the short, small diameter nerves appear to be spared) in the peripheral nerves. OPIDN, as its name suggests, has a delayed onset of approximately 1–3 weeks following an acute life-threatening exposure to an organophosphate capable of causing delayed neuropathy. The initial complaints of OPIDN include cramping of the calves with numbness and tingling in the feet and sometimes in the hands in severe cases. Next, weakness develops in the lower limbs. Bilateral foot drop and wrist drop may develop, and there are usually absent or normal reflexes. A high-stepping gait has also been described in individuals with OPIDN. There also may be motor weakness involving the limbs and motor nerve conduction studies may show abnormalities.

The current theory as to the cause of OPIDN is a two-step process that occurs in the nervous system. First, it is thought that phosphorylation of a target protein in the nervous system is required. This enzyme is known as neuropathy target esterase (NTE), formerly known as neurotoxic esterase. The biological action of NTE in the body is not known. The second, and essential, step leading to OPIDN is thought to be the transformation, or "aging" of the enzyme. This "aging" process involves cleavage of a R-group from phosphorous, resulting in a negatively charged residue attached to the active site of the enzyme. It appears that compounds that are capable of inhibiting NTE and aging can only cause OPIDN if a threshold of inhibition is reached. A high level of inhibition—70 to 80% inhibition of NTE in the brain, spinal cord, or peripheral nerve of the experimental animal—soon after dosing with an organophosphate capable of causing OPIDN is necessary before this condition can develop. Thus, the determining factor in the development of OPIDN is the formation of a critical mass of aged-inhibited NTE.

A term used to express the concentration of a substrate (such as an organophosphate compound) that is needed to inhibit 50% of an enzyme is known as the IC_{50}. One way to predict whether or not a compound will produce OPIDN is to compare AChE and NTE IC_{50}s for a specific compound. The *in vitro* and *in vivo* IC_{50}s for NTE and AChE in humans and hens (the test species used to evaluate the delayed neuropathic potential for organophosphate insecticides) for several organophosphate compounds have been compared, and it was found that for organophosphate compounds with a IC_{50} AChE/IC_{50} NTE ratio of less than one, OPIDN can only occur after recovery and treatment from acute, otherwise fatal, cholinergic crisis.

The scientific literature indicates that for at least the few compounds known to cause OPIDN that have been analyzed, the AChE IC_{50}/NTE IC_{50} typically is less than unity. This means that the concentration of a chemical that will inhibit 50% of the AChE molecules is less than the concentration of the same chemical that is required to inhibit 50% of the NTE molecules. Therefore, at a given concentration, AChE will be inhibited to a greater degree than NTE. As previously stated, it is currently theorized that more than 50% of NTE (i.e., 70–80%) must be inhibited in order to develop OPIDN. Likewise, a 50–80% inhibition of AChE results in clinical manifestations. In fact, human case reports indicate that virtually all patients who develop OPIDN were managed for a cholinergic crisis first.

It important to note that carbamates do not caused delayed neuropathy. While some carbamates are capable of inhibiting NTE, aging does not occur. In fact, experimental evidence has showed that some carbamates that inhibit NTE actually protect hens against developing OPIDN.

Biological Monitoring for Organophosphates and Carbamates

The organophosphates and carbamates have the ability to inhibit pseudocholinesterase, red blood cell cholinesterase, and nervous system cholinesterase, with biological effects being due to the actual inhibition of nervous system cholinesterase only. The levels of cholinesterase present in the blood, especially in the red blood cells, can be used to estimate the degree the nervous system is being affected by anticholinesterases. As mentioned earlier, a 50% depression of plasma and red blood cell cholinesterase levels is typically necessary before clinical manifestations are seen. Acute overexposures to organophosphates can be classified as mild (20–50% of baseline cholinesterase levels), moderate (10–20% of baseline cholinesterase levels), or severe (10% or less of baseline cholinesterase level).

Plasma cholinesterase, while susceptible to the inhibitory actions of organophosphate insecticides, has no known biological use in the body. Plasma cholinesterase can vary in an individual based on a number of disease states or conditions (e.g., decreased plasma cholinesterase levels in liver disease such as cirrhosis and hepatitis, multiple metastases, during pregnancy). Plasma cholinesterase is produced by the liver, and this enzyme is found in the nervous tissue, heart, pancreas, and in white matter of the brain. Plasma cholinesterase levels typically decline and regenerate more rapidly than red blood cell cholinesterase levels. Plasma cholinesterase levels typically regenerate at the rate of 25% in the first 7–10 days. Following organophosphate intoxication, plasma cholinesterase levels may remain depressed for a period of 1–3 weeks.

The enzyme in red blood cell cholinesterase is the same enzyme that is present in the nervous system. Red blood cell cholinesterase regenerates at the rate of approximately 1% per day in the body and is dependent on the synthesis of new red blood cells in the body. As mentioned earlier, in severe intoxications from organophosphate pesticide exposure, red blood cell cholinesterase could take as long as 3 months to regenerate.

The measurement of cholinesterase activity in cases of carbamate intoxication is not useful due to quick reactivation of cholinesterase following carbamate overexposures.

Treatment for Organophosphate and Carbamate Intoxication Symptomatology

Each year, approximately 3 million individuals suffer from OP intoxication worldwide, resulting in about 300,000 deaths. Historically, two treatments have been used in cases of organophosphate intoxication: (i) atropine (muscarinic antagonist) and (ii) an oxime medication such as pralidoxime or 2-PAM (reactivates acetylcholinesterase). Atropine competes with muscarinic sites, and treatment with the agent ameliorates the symptoms of nausea, vomiting, abdominal cramps, sweating, salivation, and miosis. Atropine treatment has no effect on the nicotinic signs, such as muscle fasciculations, muscle weakness, or respiratory failure. Atropine does not reactivate cholinesterase.

The second therapeutic agent, 2-PAM, is a medication that reactivates the organophosphate-inhibited cholinesterase enzyme by the removal of the phosphate group that is bound to the esteratic site. 2-PAM should be given fairly soon after exposure due to the fact that the aged enzyme cannot be reactivated. 2-PAM is effective in improving the symptoms of respiratory depression and muscle weakness. Apparently, there is controversy regarding the usefulness of oximes in treating OP poisoning. In 2011, the Cochrane Group conducted an evidence-based evaluation of whether the oximes are a useful treatment in OP pesticide intoxication. The Cochrane Group concluded, based on their evaluation of the literature, that *current evidence is insufficient to indicate whether oximes are harmful or beneficial.* In another recent review published in 2011 in the *British Medical Journal*'s *Clinical Evidence*, oximes were classified as having *unknown effectiveness* in treating individuals suffering from acute OP poisoning. Individuals suffering from carbamate intoxication should not be treated with 2-PAM possibly because the reversal of the carbamate inhibitor could add insult to injury.

Decontamination of the individual should include the removal of any contaminated clothing (rubber gloves should be worn to avoid contact with contaminated clothing and materials) and thorough washing of the contaminated skin with soap and water. Other treatments include resuscitation, oxygen and/or respiratory support, and/or gastric decontamination, if necessary. No specific treatment for OPIDN exists; GABA antagonists may be used if spasticity occurs.

Regulatory Information on Organophosphates and Carbamates

OSHA PELs and ACGIH TLVs exist for some of the organophosphate and carbamate insecticides. Biological exposure indices (BEIs) also exist for exposure to AChE-inhibiting pesticides as well as parathion (see Table 17.4). There is an ACGIH BEI *p*-nitrophenol levels (metabolite of parathion) in urine as well as a BEI for cholinesterase activity for workers exposed to organophosphate cholinesterase inhibitors.

TABLE 17.4 2012 ACGIH Biological Exposure Indices (BEIs) for Various Pesticides

Pesticide Determinant	Sampling Time	BEI
Acetylcholinesterase inhibiting pesticides	Discretionary	70% of individual's baseline
Cholinesterase activity in red blood cells		
Parathion		
Total *p*-nitrophenol in urine	End of shift	0.5 mg/g creatinine
Cholinesterase activity in red cells	Discretionary	70% of individual's baseline
Pentachlorophenol		
Total PCP in urine	Prior to last shift of workweek	2 mg/g creatinine
Free PCP in plasma	End of shift	5 mg/l

17.2 ORGANOCHLORINE INSECTICIDES

Introduction, Use, and History of Organophosphates and Carbamates

While organochlorine insecticides had widespread use in the 1940s through the mid-1960s in agricultural and malarial control programs, their use has become almost completely discontinued due concern over environment effects. Examples of organochlorine insecticides that were commonly used in the past include toxaphene (Toxakil™), endrin (Hexadrin™), aldrin (Aldrite™), endosulfan (Thiodan™), BHC (hexachlorocyclo-hexane), dienoclor (Pentac™), heptachlor (Heptagran™), dicofol, mirex (Declorane™), chlordane, lindane, and DDT. One organochlorine compound that is still in use today is lindane, which is used in the medicinal product Kwell® for human ectoparasite disease as a 1% lotion or shampoo product (as approved by the U.S. Food and Drug Administration). In addition to their use as insecticides in agricultural and forestry settings, organochlorine compounds were also previously widely used as structural protection against termites in the past.

Physical/Chemical Properties

The physical and chemical properties of the organochlorine compounds, that is, their lipophilicity, low vapor pressures, and slow rate of degradation, made them not only effective pesticides, but these same qualities also resulted in their persistence in the environment and their bioaccumulation in the food chain leading to the eventual discontinuance of their use.

Mechanisms of Action

Organochlorine compounds chemicals, at sufficient doses, can act on the nervous system to produce adverse effects. This class of chemicals is thought to act by the interference with cation exchange across the nerve cell membranes, resulting in hyperactivity of the nerves.

Benzene hexachloride compounds (BHC) (lindane and related compounds) are examples of isomers that produce different effects on the nervous system. The γ isomer, also referred to as lindane, causes severe convulsions with rapid onset, while other isomers of BHC generally cause central nervous system depression. The relative contribution of each of the isomers may explain toxicological differences between formulations of these products. The effect of the cyclodiene organochlorine compounds, for example, dieldrin, is on the central nervous system.

Pharmacokinetics

Organochlorine compounds are lipophilic and can be absorbed not only through the intestines, but also across the lung and skin. Some of these compounds, for example, lindane, endrin, and chlordane, are more readily absorbed dermally than other compounds in this class, such as DDT or toxaphene.

This class of insecticides, for example, DDT, can be also be stored in fatty tissues in the body. Organochlorine compounds can be detected in adipose tissue, serum, and in milk. Some compounds, for example, DDT, are mainly stored unchanged in adipose tissue (some DDE is stored in adipose tissue), while other, for example, endrin, are stored in a metabolized form, aldrin metabolized to endrin (this transformation also occurs in the environment) or heptachlor metabolized to heptachlor epoxide. Organochlorine insecticides are eliminated primarily via the feces.

Acute and Chronic Health Effects of Organochlorine Insecticides

The principal adverse effect associated with overexposure to organochlorine insecticides is nervous system hyperactivity (e.g., headache, dizziness, paresthesias, tremor, incoordination, or convulsions). Early symptoms seen in chlorinated insecticide intoxications, such as with DDT, include hyperesthesias and paresthesias of the face and limbs, dizziness, nausea and vomiting, headache, tremor, and mental disturbances. Myoclonic movement and convulsions are sometimes seen in severe cases of poisoning. It should be noted that with overexposure to the toxaphene and cyclodiene compounds (e.g., aldrin, endrin, chlordane, and heptachlor) the first sign seen is convulsions, in the absence of the early symptoms just mentioned. Convulsions seen in these cases may not first occur until 2 days after exposure. A group of factory workers overexposed to chlordecone (Kepone™) manifested signs and symptoms, including gait disturbances, opsoclonus, headache, tremors, hepatomegaly/splenomegaly, and neurobehavioral changes.

While studies of the carcinogenicity of organochlorine compounds have demonstrated positive effects in mice, but generally not in rats, at high doses, there is generally no evidence of cancer in humans, even in the most highly exposed individuals, for example, workers involved in the manufacture and formulation of organochlorine insecticides.

Biological Monitoring for Organochlorine Insecticides

Levels of organochlorine insecticides can be detected at background levels in biological tissues of individuals not occupationally exposed to these compounds. Serum and adipose tissue testing for the presence of organochlorine pesticides in the general population has been conducted. The National Human Adipose Tissue (NHAT) survey was conducted in 1982. In this study, 763 individual adipose tissue specimens collected from the general population were tested for various compounds, including several organochlorine compounds (β-BHC, *p,p′*-DDE, dieldrin, heptachlor

epoxide, and DDT). These results are presented in the NHATS Broad Scan Analysis. Results for organochlorine insecticides detected in the serum of the general population are reported in Health and Nutrition Examination Survey II (HANES II).

Treatment of Organochlorine Intoxication

Treatment of organochlorine intoxication is supportive (e.g., control of convulsions with benzodiazepines or barbiturates). One chlorinated insecticide that has been shown that can be effectively removed from the body is chlordecone (Kepone, the compound involved in the poisoning of factory workers discussed earlier). In this case, chloestyramine was used therapeutically to treat the workers who had been poisoned with chlordecone. In nine patients, administration of 24 g of chloestyramine per day resulted in a 3.3- to 17.8-fold increase in fecal elimination of chlordecone. Treatment also resulted in a reduction of chlordecone half-life in blood from 165 to 80 days.

17.3 INSECTICIDES OF BIOLOGICAL ORIGIN

Many compounds are present in nature that have insecticidal qualities, including extracts from the chrysanthemum flower and from the *Legumionocae* genera (e.g., rotenone). Trade names of insecticides in this classification include Pyrocide™ (pyrethrum) and Prentox™ (rotenone).

Pyrethrum and Pyrethrins

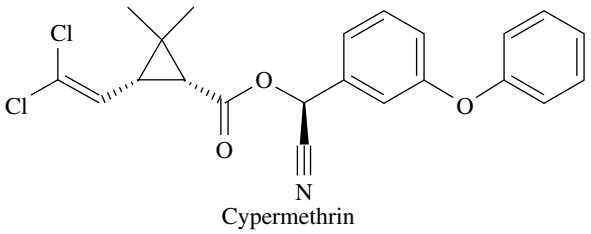
Pyrethrum is an extract from the chrysanthemum flower, *Chrysanthemum cinerariaefollium* ("Dalmatian insect flowers") and other species. This extract contains approximately 50% natural pyrethrins—the insecticidal component of the extract. The USEPA states the term *pyrethrin* refers to all of the six isomers that are found the chrysanthemum flower (jasmolin 1 and 2; cinerin 1 and 2; pyrethrins 1 and 2). These compounds are often formulated with a synergist such as piperonyl butoxide or *n*-octyl bicycloheptene dicarboximide. These synergists are incorporated in order to slow down the degradation of the pyrethrin compounds. Pyrethrins have been registered as insecticides since the 1950s, and there are over 1300 pyrethrin-containing pesticide products registered currently. This class of compounds are commonly used in household insecticides and in pet products (e.g., flea and tick dips and sprays). Pyrethrin is used in lice medications also and was used as an antihelmintic in the past.

Pyrethrins and pyrethrum are very rapidly metabolized and excreted from humans and have very low mammalian toxicity. Crude pyrethrums have been associated with allergic-type responses in individuals, although this action is most likely due to the noninsecticidal components of this compound. The USEPA has reported that pyrethrins have low toxicity and are not expected to have any neurological effects on humans. Treatment of pyrethrin and pyrethrum exposure is primarily symptomatic.

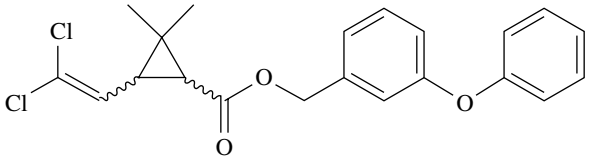
Synthetic Pyrethroids

Because pyrethrins and pyrethrum insecticides are unstable in light and heat, synthetic pyrethroids, which have better stability to light and heat, were developed for use in agricultural settings as well as for home pest control. Over 1000 synthetic pyrethroids have been developed over the years and include compounds such as bifenthrin, cyfluthrin, cypermethrin (Cymbush™), deltamethrin, fenpropathrin (Danitol™), fluvalinate (Mavrik™), permethrin (Ambush™), resmethrin, and tralomethrin (Scout™) (Figure 17.3).

The site of action of the pyrethroids are the voltage-dependent sodium channels in nerves. The general basis for nerve impulse generation and conduction is the ionic permeability of the membrane combined with the sodium (high levels outside the cell) and potassium (high levels inside in the cell) concentration gradients. The resting cell membrane is maintained by the sodium-potassium pump, and the inside of the cell is negative respective to the outside of the cell. A normal nerve impulse is caused by a quick transient increase in the permeability of the membrane to sodium ions, causing an inward influx of sodium, followed by an increase in the potassium permeability causing an outward flow of potassium. The ionic currents cause a temporary reversal of the membrane potential from negative to positive, resulting in nerve impulse conduction along the nerve fiber. Pyrethroids exert their effect by slowing the closing of the sodium activation gate. Type I pyrethroids prolong individual channel currents causing whole cell sodium influx to be prolonged elevating the after-potential until the threshold potential is reached and repetitive discharges occur. Examples of



Cypermethrin



Permethrin

FIGURE 17.3 Synthetic pyrethroids.

type I pyrethroids include allethrin, cismethrin, permethrin, and resmethrin. Type I pyrethroids, at high levels in animals, have been reported to cause increased sensitivity to external stimuli, tremors, increased body temperature, and rigor immediately preceding death. Type II pyrethroids cause an extremely prolonged sodium current, leading to depolarization of the nerve and impulse conduction block. Type II pyrethroids include cyfluthrin, cyhalothrin, cypermethrin, deltamethrin, and fenvalerate. Type II pyrethroids cause behavioral problems early in the intoxication, leading to salivation, miosis, bradycardia, tremor, decreased startle response to sound, and ataxia.

Like pyrethrins and pyrethrum, the synthetic pyrethroids are rapidly metabolized and excreted in humans and do not bioaccumulate. In fact, the relatively resistant nature of mammals, including man, to the ability to metabolize these compounds quickly and efficiently. Synthetic pyrethroids have greater insecticidal activity and lower mammalian toxicity than the organophosphate, carbamate, and organochlorine insecticides. Experimental animals that have been treated with high doses of pyrethroids tremors, salivation, and/or convulsions. In general, animals surviving an acute intoxication to pyrethroids recover within several hours of exposure.

The primary reported reaction to exposure to synthetic pyrethroid insecticides in humans occurs with exposure to those pyrethroids containing cyano-groups (e.g., fenvalerate and cypermethrin). This reaction consists of paresthesia, typically occurring around the mouth region in workers exposed to these compounds. This paresthesia is reversible and dissipates usually within 24 h of cessation of exposure. An occupational study of 199 workers who were involved in dividing and packaging pyrethroids (fenvalerate, deltamethrin, and cypermethrin) showed that aside from transient paresthesias occurring in the facial area and sneezing and increased nasal secretions, there were essentially no adverse health effects attributable to the pyrethroid exposure. Thus, while acute exposure to high concentrations of pyrethroids has been reported to cause nervous system symptoms in some workers, it appeared to recovery does occur. Treatment of pyrethrin overexposure consists of decontamination and supportive type treatment.

As an aside, the 2010 American Academy of Pediatrics Clinical Policy on Head Lice states that, "*Permethrin has been the most studied pediculicide in the United States and is the least toxic to humans*." and "*Permethrin is a synthetic pyrethroid with extremely low mammalian toxicity*."

Rotenone

Rotenone (Noxfish™) occurs naturally in several plants species, for example, the *Leguminocae* genera, and is used mainly as an insecticide as well as to eliminate fish (piscicide) in lakes and ponds. The mechanism of action for rotenone is as a respiratory toxin, blocking electron transport at ubiquinone, preventing oxidation of NADH. Rotenone seems to have low toxicity in man with few reports of serious injury reported. Occupational exposure to the powder of the plant that contains rotenone has reportedly caused dermal and respiratory tract irritation and numbness in mouths of workers. Recent articles in the literature have attempted to link exposure to rotenone with development of Parkinson's disease as part of the "environmental hypothesis" that exposure to chemicals in the environmental contributes to the etiology of the disease, but there appears to be a lack of a causal role of rotenone in the role of Parkinson's disease in humans. Treatment of rotenone overexposure consists mainly of decontamination and supportive therapy.

Bacillus thuringiensis Microbial insecticides, such as several strains *Bacillus thuringiensis* (e.g., Dipel™), have been developed as effective insecticides. The endotoxin of *B. thuringiensis* is insecticidal in certain sensitive species via cell membrane disruption in susceptible insect species, not infectious activity. *B. thuringiensis* has not generally been associated with mammalian or human toxicity, with only rare instances of adverse effects in humans being reported. A group of 18 human volunteers ingested 1 g of a *B. thuringiensis* formulation for 5 days, with 5 of these 18 subjects also inhaling 100 mg of the powder for 5 days with no reported adverse effects. Further, a group of workers exposed to various processes involved in the formulation of a commercial product containing the biological showed no adverse health effects.

17.4 HERBICIDES

Chlorophenoxy Herbicides

The chlorophenoxy herbicides 2,4-dichlorophenoxy acetic acid (2,4-D) and 2,4,5-trichlorophenoxy acetic acid (2,4,5-T) are probably the most commonly recognized of the chlorophenoxy herbicides. These compounds exert their action in plants by acting as growth hormones, but have no such hormonal action in animals or humans. Some examples of chlorophenoxy herbicides include Banvel® (dicamba), Weedone® (2,4-D) (Figure 17.4), Basagran M® (MCPA), and 2,4-DB (Butoxone®).

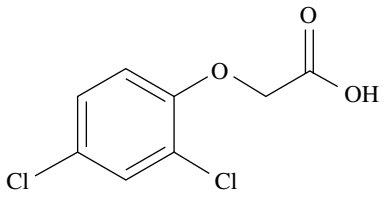


FIGURE 17.4 2,4-Dichlorophenoxyacetic acid (2,4-D).

Acute Toxicity

2,4-Dichlorophenoxyacetic acid (2,4-D) is prepared commercially by the reaction of 2,4-dichlorophenol and monochloroacetic acid. Other chlorophenoxy herbicide analogs include 2,4-DB, 2,4-DP, MCPA ((4-chloro-2-methylphenoxy) acetic acid), MCPP, and the herbicides 2,4,5-trichlorophenoxyacetic acid (2,4,5-T) and 2-(2,4,5-trichlorophenoxy) propionic acid (2,4,5-TP; Silvex™) that are no longer used.

Dioxin (2,3,7,8-TCDD) has not been identified in 2,4-D formulations (WHO, 1987). While in the past 2,3,7,8-TCDD contamination may have occurred in 2,4-D, this was due to contamination from the production of 2,4,5-T. The synthesis of 2,4-D does not produce 2,3,7,8-TCDD.

The primary routes of exposure to chlorophenoxy herbicides are dermal and inhalation. The mechanism of chlorophenoxy compounds are the uncoupling of oxidative phosphorylation and by decreasing oxygen consumption in tissue. These compounds are fairly rapidly excreted and do not accumulate in the body. These compounds are excreted via the urine primarily, and apart from conjugation of acids, little biotransformation occurs in the body.

Following ingestion, the acute toxicity of chlorophenoxy herbicides includes irritation of the mucous membranes and gastrointestinal lining. Large intentional overdoses with chlorophenoxy acids have resulted in symptoms of coma, metabolic acidosis, myotonia, mucous membrane irritation, and myalgias. While cases of peripheral neuropathy following exposure to 2,4-D have been reported sporadically throughout the literature, no causal association between this compound and neuropathy has been proven. Treatment of cases of overexposure with chlorophenoxy herbicides is symptomatic and also involves decontamination.

Carcinogenicity

2,4-D is currently classified as a "D" carcinogen (not classifiable) by the USEPA's Office and Pesticide Programs. MCPA is classified as *Not Likely to be a Human Carcinogen.* A recent review of the available animal and human data for the chlorophenoxy herbicides 4-chloro-2-methyl phenoxyacetic acid (MCPA), 2-(4-chloro-2 methylphenoxy) propionic acid (MCPP), and 2-(2,4-dichlorophenoxy) propionic acid (2,4-DP) and concluded that there was no evidence to indicate that these compounds were carcinogenic to humans.

Bipyridyl Compounds: Paraquat and Diquat

Paraquat (1,1′-dimethyl-4,4′-dipyridylium) (Figure 17.5) and diquat (1,1′-ethylene-2,2′-bipyridylium) are bipyridylium herbicides, with common trade names including Gramoxone® (paraquat) and Reglone®, Weedtrine®, or Reward® (diquat). Due to its acute toxicity, paraquat is classified by the USEPA as a Restricted Use Pesticide. A majority of reported cases of toxicity associated with

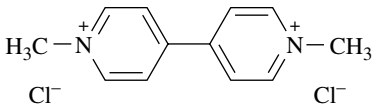


FIGURE 17.5 Paraquat.

both paraquat and diquat are seen in cases of accidental or intentional (suicidal) ingestion, with paraquat having greater toxicity than diquat. Valeric acid is added to paraquat solutions to act as an emetic and stenching agent. Paraquat solutions may have a blue dye added to them in order to keep it from being confused with being a beverage and being ingested.

Paraquat poisoning (e.g., from suicide attempt) can lead to multiorgan toxicity (e.g., gastrointestinal tract, kidney, heart, and liver), coma, and pulmonary fibrosis. Early deaths occurring after intoxication with paraquat result from acute pulmonary edema, oliguric renal failure, and hepatic failure. Deaths occurring 1–3 weeks following an intoxication episode are typically the result of pulmonary fibrosis. Paraquat is not typically readily dermally absorbed, but reports of toxicity following sufficient dermal absorption have been seen in individuals with skin abrasions or a rash or individuals with continued dermal exposure to paraquat. Sufficient dermal exposure to paraquat can also cause dermal irritation, blistering, and ulceration. Similar irritant effects are seen in the esophagus and stomach of individuals swallowing paraquat. Paraquat concentrates in the lung, where its proposed mechanism of action leading to pulmonary fibrosis is free radicals being generated leading to lipid peroxidation. Pulmonary fibrosis, which can be fatal in cases with sufficient exposure, begins within 2 days to 2 weeks following paraquat exposure.

Paraquat is considered to be a nonvolatile compound, making the inhalation route an unlikely route of exposure. Aerosol paraquat droplets have been measured as having diameters exceeding 5 μm, indicating that they do not reach the alveolar membrane to cause either direct or systemic toxicity via inhalation. In two field trials in which absorption of paraquat was measured by urinary paraquat levels, systemic absorption was apparently not significant. The authors of that study concluded that ordinary care in personal hygiene is sufficient to prevent any hazard from surface injury or from systemic absorption. Another study conducted on a group of 85 paraquat spray-men revealed no adverse health effects (aside from irritant-type effects), including no lung effects, attributable to long-term occupational use of this herbicide. Similar to rotenone, there has been some suggestion in the literature that implicates paraquat in the pathogenesis of Parkinson's disease in humans; however, reviews have found that the evidence of such a causal relationship to be is limited and insufficient.

Diquat causes less dermal irritation and injury than does paraquat, and diquat is not selectively concentrated in

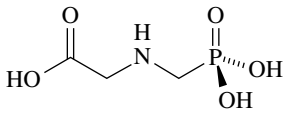


FIGURE 17.6 Glyphosate.

pulmonary tissue as paraquat. Diquat, in contrast to paraquat, causes little to no injury to the lungs; however, diquat has an effect on the CNS, whereas paraquat does not. The mechanism of action of diquat is thought to be similar to that of paraquat, involving the production of superoxide radicals that cause lipid membrane destruction. Dermal exposure to sufficient levels of diquat can cause fingernail damage, irritation of the eyes and mucous membranes. Intoxication by diquat via the oral route has reportedly caused signs and symptoms including gastrointestinal irritation, nausea, vomiting, and diarrhea. Both paraquat and diquat are reportedly associated with renal toxicity. There is no known specific antidote for either paraquat or diquat poisoning.

Glyphosate (Round-Up®) (*N*-(phosphonomethyl) glycine) (Figure 17.6) is a herbicide that interferes with amino acid metabolism in plants. In animals, it is thought to act as a weak uncoupler of oxidative phosphorylation. Glyphosate is moderately absorbed through the gastrointestinal tract, undergoes minimal biotransformation, and is excreted via the kidneys. There have been several reports in the literature of intoxications, typically resulting from accidental or suicidal ingestion, following overexposed to the glyphosate-containing product Round-Up. Various signs and symptoms including gastrointestinal irritation and damage, as well as dysfunction in several organ systems (e.g., lung, liver, kidney, CNS, and cardiovascular system). It has been proposed that the toxicity seen following intoxication with Round-Up is due to the surfactant agent in the commercial product, polyethoxylated tallow amine (POEA). One study conducted determined that the irritancy potential of the commercial preparation of Round-Up is similar to that of baby shampoo. Glyphosate is a USEPA Group E carcinogen (Evidence of Non-Carcinogenicity to Humans). A fairly recent review found that glyphosate was nonmutagenic, noncarcinogenic, and nonteratogenic and causes no effects on reproduction. Glyphosate did not show evidence of bioaccumulation. In addition, this same reviewed concluded that under normal conditions of use, Round-Up herbicide is not expected to cause any health risks to humans and that it does not cause any harmful effects on development, reproduction, or the endocrine system.

Symmetrical Triazines

Examples of triazine and triazole herbicides include atrazine (2-chloro-4-ethylamino-6-isoproplyamine-*s*-triazine), propazine, simazine (2-chloro-4,6-bis(ethylamino)-*s*-triazine), and cyanazine (2-chloro-4-(1-cyano-1-methylethylamino)-6-ethylamino-*s*-triazine). Triazine herbicides have relatively low toxicity, and no cases of systemic poisoning have appeared to have been reported. Occasional reports of dermal irritation from exposure to triazine herbicides have occurred in the literature. Some of the triazine herbicides have been ranked by USEPA's Office of Pesticide Program: (i) atrazine: Not Likely to be Carcinogenic to Humans; (ii) propazine: Not Likely to be Carcinogenic to Humans; and (iii) simazine: Not Likely to be Carcinogenic to Humans. While atrazine has been reported to cause mammary tumors in Sprague–Dawley rats, a fairly recent review and risk assessment concluded that this lone finding has no relevance for humans (i.e., the finding in female Sprague–Dawley rats is a strain-, sex-, and tissue-specific effect). This conclusion was also similar to that noted in a 2011 review on this issue by researchers at the University of Texas—*Epidemiological studies provide no support for a causal relationship between atrazine exposure and breast cancer. This conclusion is consistent with IARC's classification of atrazine as "unclassifiable as to carcinogenicity" and the USEPA's classification of atrazine as "not likely to be carcinogenic.*"

17.5 FUNGICIDES

Fungicides are compounds that are used to control the growth of fungi and have found uses in many different products, from their use to protect grains after harvesting while they are in storage, and in paint.

Pentachlorophenol

Pentachlorophenol, also known as penta, is used as a wood preservative for fungus decay or against termites as well as a molluscicide. Trade names of pentachlorophenol (Figure 17.7) include Pentacon, Penwar™, and Penchlorol™.

Pentachlorophenol is readily absorbed via the skin, lung, and gastrointestinal tract. Pentachlorophenol and its biotransformation products are excreted primarily via the kidneys. The biochemical mechanism of action of pentachlorophenol

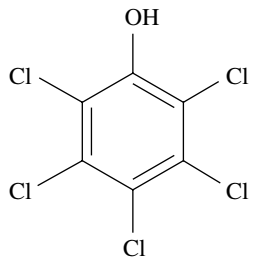


FIGURE 17.7 Pentachlorophenol

is through increase in oxidative metabolism from the uncoupling of oxidative phosphorylation. This increase in oxidative metabolism in poisonings can lead to an increase in body temperature. In fatal cases of poisoning from pentachlorophenol, body temperature as high as almost 41.8 °C (107.4 °F) have been reported. Severe overexposure to pentachlorophenol can cause signs and symptoms including delirium, flushing, pyrexia, diaphoresis, tachypnea, abdominal pain, nausea, and tachycardia.

Because pentachlorophenol volatilizes from treated wood and fabric, excessively treated indoor surfaces can lead to irritation of the skin, eyes, and upper respiratory tract. Contact dermatitis has been reported in workers exposed dermally to pentachlorophenol. Treatment of pentachlorophenol poisoning consists mainly of decontamination of clothing and skin and/or gastrointestinal tract as well as supportive treatment for symptoms associated with the exposure (e.g., temperature control).

Pentachlorophenol can be assayed for in blood, urine, and adipose tissue. The ACGIH biological exposure index for pentachlorophenol is 2 mg/g creatinine total pentachlorophenol in urine prior to the last shift of the workweek or 5 mg/l free pentachlorophenol in plasma at the end of the work shift.

Dithiocarbamates/Thiocarbamates

The dithiocarbamates and the thiocarbamates have little insecticidal toxicity, unlike the *N*-methyl carbamates (e.g., the acetylcholinesterase-inhibiting carbamate, carbaryl) discussed earlier, and are used as fungicidal compounds. Examples of thiocarbamate fungicides include thiram (Aules®), metamsodium (Metam®, Vapam®), ziram (Vancide®, Ziram 76®), ferbam (Fermbam®), and the ethylene *bis* dithiocarbamate (EBDC) compounds—maneb (Manex®), zineb, metiram (Polyram®), nabam (Aquatreat®), and mancozeb (Rainshield®, Dithane®, Pennocozeb®).

Thiram dust has been reported to cause eye, skin, and mucous membrane irritation, with contact dermatitis and sensitization reportedly occurring in a few workers. Systemic intoxications that have been associated with exposure to thiram have resulted in symptomatology similar to that caused by reactions to disulfiram (Antabuse), a dithiocarbamate medication that is used to treat alcoholism. Thiram, like disulfiram, is not a cholinesterase inhibitor, but does cause inhibition of the enzyme acetaldehyde dehydrogenase (the enzyme responsible for the conversion of acetaldehyde to acetic acid), and reportedly, in rare cases, workers who have been exposed to thiram have complained of "Antabuse" reactions after ingestion of alcoholic beverages.

Exposure to ziram, ferbam, and the EBDC compounds has been associated with skin, eye, and respiratory tract irritation in humans. Maneb and zineb have been associated with cases of chronic dermatological disease, possibly due to dermal sensitization to these compounds in workers. It is reported that systemic illness in humans has been rare.

Chlorothalonil

Chlorothalonil (Bravo™, Daconil™) (2,4,5,6-tetrachloro-1,3-benzenedicarbo-nitrile), a fungicide and mildewicide, has been reported to cause dermal and mucous membrane irritant effects in humans exposed to this compound. Chlorothalonil appears to have low potential for toxicity in humans. While the USEPA's Office of Pesticide Program has categorized chlorothalonil as *Likely to be Carcinogenic to Humans*, they have concluded that it is threshold carcinogen in animals, and a recent epidemiological study in humans found no strong evidence for cancer among pesticide applicators exposed to chlorothalonil.

Copper Compounds

Copper compounds are used as fungicides as well as antimicrobials. Exposure to dust and powder formulations of copper-based fungicides has been reported to cause irritation of the skin, eyes, and respiratory tract. Systemic intoxication in humans by copper fungicides has been rarely reported. Ingestion of the compound has reportedly caused gastrointestinal irritation, nausea vomiting, and diarrhea, headache, sweating, weakness, liver enlargement, hemolysis and methemoglobinemia, albuminuria, hemoglobinurina, and occasionally renal failure. Treatment of copper intoxication can include an effort to prevent absorption (e.g., lavage) following by chelation therapy.

17.6 RODENTICIDES

The rodenticides, as the name indicates, are a class of compounds designed to specifically target rodents. These compounds have, in some cases, taken advantage of physiological differences between rodents and other mammals (namely human) that makes rodents more susceptible to their toxic effects. The most efficient route of exposure of these compounds for rodents is via ingestion.

Coumarins and Indandiones

This class of rodenticides work by depression of the vitamin K synthesis of the blood clotting factors II (prothrombin), VII, IX, and X. This anticoagulant property manifests as diffuse internal hemorrhaging occurring typically several days of rodenticide bait ingestion. Warfarin (Figure 17.8) is a commonly used coumarin rodenticide that causes its toxic effects by inhibiting the formation of prothrombin and the inhibition of vitamin K-dependent factors in the body. Other anticoagulant rodenticides include coumafuryl, brodifacoum,

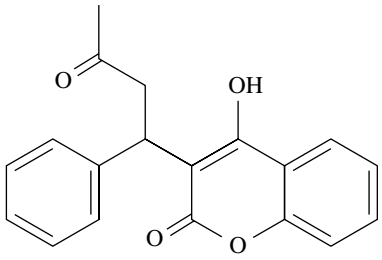


FIGURE 17.8 Warfarin

difenacoum, and prolin. Warfarin is known to be absorbed both dermally and from ingestion. Signs and symptoms of intoxication with warfarin include epistaxis, hemoptysis, bleeding gums, gastrointestinal tract and genitourinary tract hemorrhage, and ecchymoses.

The indandiones, unlike the coumarins, cause nervous system, cardiac, and pulmonary effects in laboratory animals preceding the death from the anticoagulant effects. These types of adverse effects have not been reported in cases of human exposure. Examples of indandione rodenticides include diphacinone, diphacin, and chlorphacinone.

The most prominent clinical laboratory sign from the administration of these classes of compounds is an increased prothrombin time and a decrease in plasma prothrombin concentration. Treatment of toxicity from coumarins and indandiones consists of the administration of vitamin K1.

Thallium Sulfate

Thallium sulfate is readily absorbed via ingestion and dermally, as well as via inhalation. The target organs of thallium sulfate include the gastrointestinal tract (hemorrhagic gastroenteritis), heart and blood vessels, kidneys, liver, skin, and the hair. Symptoms such as headache, lethargy, muscle weakness, numbness, tremor, ataxia, myoclonia, convulsions, delirium, and coma are seen in cases of thallium sulfate–induced encephalopathy. Death from thallium sulfate intoxication is due to respiratory paralysis or cardiovascular failure.

Serum, urine, and hair thallium levels can be used to assess exposure to this compound. There is no specific treatment for thallium sulfate poisoning, and treatment is supportive. Syrup of ipecac and activated charcoal can be used to decrease gastrointestinal absorption.

Sodium Fluoroacetate (1080®)

Sodium fluoroacetate, also known as 1080. Not only is sodium fluoroacete used as a rodenticide, but it is has been injected into reservoirs of collars (“toxic collars”) that are attached to the necks of goats and sheep to prevent against coyotes that might prey on these domestic farm animals. When the coyote bites the neck breaking the collar, it is likely that sodium fluoroacetate will be released, thus poisoning the prey. Due to its toxicity, sodium fluoroacetate is an USEPA Restricted Use Pesticide. This compound is easily absorbed via ingestion as well as through inhalation and dermal routes. Sodium fluoroacetate’s toxicity is due to the reaction of three molecules of fluoroacetate that forms fluorocitrate in the liver. Fluorocitrate adversely affects cellular respiration through disruption of the tricarboxylic acid cycle (inhibiting the enzyme *cis*-aconitase). It is thought that the accumulation of citrate in tissues also accounts for some of the acute toxicity associated with this compound. The target organs of sodium fluoroacetate are the heart (seen as arrhythmias leading to ventricular fibrillation) and the brain (manifested as convulsions and spasms), following intoxication (typically following suicidal or accidental ingestion). A specific antidote to sodium fluoroacetate intoxication does not exist. Treatment consists of decontamination and supportive therapy, including gastric lavage and catharsis.

17.7 FUMIGANTS

The fumigants are a group of compounds that are volatile in nature. Some of the fumigants exist in a gas phase at room temperature while others are liquids or solids. Fumigants are in general readily absorbed via dermal, respiratory, and ingestion routes. Treatment for overexposure to fumigants typically includes irrigation of the contaminated areas (skin, eyes). Following irrigation of eyes, medical treatment should be sought due to the fact that some of these compounds are severely corrosive to the cornea. Sufficient dermal absorption may occur as to produce systemic effects. Patients with inhalation exposure should be monitored for pulmonary edema and treated accordingly if edema develops. Contaminated clothing should be removed and discarded. It should be noted that certain fumigants have the ability to penetrate rubber and neoprene (often used for personnel protective equipment).

Methyl Bromide

Methyl bromide (Brom-O-Sol™, Terr-O-Gas™) (Figure 17.9) has been in use as a fumigant since 1932 and is a colorless and practically odorless gas, with its low warning potential

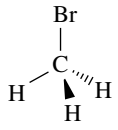


FIGURE 17.9 Chemical structure of methyl bromide.

contributing to its toxicity. Chloropicrin is sometimes added as a warning agent. At higher concentrations, the odor of methyl bromide is similar to chloroform. Fatalities have been reported during application and from early re-entry into treated areas. Methyl bromide has been used to treat dry packaged foods in mills and warehouses as well as used as a soil fumigant to control nematodes and fungi.

Methyl bromide is very irritating to the lower respiratory tract. It is thought that the parent compound is responsible for the toxicity of the methyl bromide, with the mechanism of toxicity possibly having to do with its ability to methylate sulfhydryl enzymes. Exposure to high concentrations of methyl bromide can lead to pulmonary edema or hemorrhage, and typically of delayed onset (several hours after exposure). Symptoms of acute intoxication include those consistent with central nervous system depression such as headache, dizziness, throat irritation, nausea, cough visual disturbances, vomiting, and ataxia. Exposures to very high concentrations can lead to unconsciousness. In cases of exposure to fatal levels of methyl bromide, death typically occurs within 4–6 h to 1–2 days postexposure, with the cause of death being respiratory or cardiovascular failure resulting from pulmonary edema. Dermal exposure to liquid methyl bromide can cause skin damage in the form of burning, itching, and blistering. Treatment of methyl bromide poisoning is symptomatic and also includes removal of contaminated clothing. The USEPA Office of Pesticide Programs has classified methyl bromide as *Not Likely to be Carcinogenic in Humans*.

Ethylene Oxide

Ethylene oxide, also known as epoxyethane or ETO (Figure 17.10), is used in chemical manufacturing as well as for sterilizing medical equipment and fumigating spices. USEPA has determined that ethylene oxide would be effective against *Bacillus anthracis* spores. Ethylene oxide exists as a colorless gas and that has a high odor threshold. Ethylene oxide also is a severe mucous membrane and skin irritant. Dermal exposure at sufficient levels can result in edema, burns, blisters, and frostbite. Acute intoxications can result in central nervous system depression characterized by headache, nausea, vomiting, drowsiness, weakness, and cough. Exposure to extreme concentrations of ethylene oxide can cause the development of pulmonary edema and cardiac arrhythmias. The International Agency for Research on Cancer recently classified ethylene oxide as a Group 1 carcinogen (Carcinogenic to Humans) based on evidence that it is an alkylating agent that directly reacts with DNA; IARC concluded there was only limited evidence in humans for the carcinogenicity of ethylene oxide. It should be noted that while ethylene oxide is a fumigant, it is also a compound that is formed endogenously (in our bodies).

FIGURE 17.10 Chemical structure of ethylene oxide.

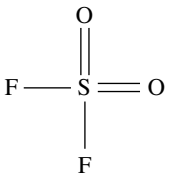


FIGURE 17.11 Chemical structure of sulfuryl fluroide.

Sulfuryl Fluoride

Sulfuryl fluoride (Vikane™) (SO_2F_2) (Figure 17.11), a colorless and odorless gas, is used as a structural fumigation. Fatalities have been reported from individuals entering buildings recently fumigated with sulfuryl fluoride before re-entry was allowed. The acute toxic effects from sulfuryl poisoning include mucous membrane irritation, nausea, vomiting, dyspnea, cough, severe weakness, restlessness, and seizures. The USEPA has classified sulfuryl fluoride as *Not Likely to be Carcinogenic to Humans.*

17.8 SUMMARY

This chapter has discussed the toxicology of some of the most commonly used groups of pesticides:

- Organophosphate and carbamate insecticides
- Organochlorine insecticides
- Insecticides of biological origin
- Herbicides
- Fungicides
- Rodenticides
- Fumigants

From the discussion included in this chapter, the following are the main points to be gained:

- Pesticides are used for a variety of different reasons including control or eradication of pests from homes, pets, or crops. Pesticides are also important in the control of vector-borne diseases (e.g., malaria, West Nile Virus).
- Individuals may be exposed to a variety of pesticides via inhalation, ingestion, or dermal routes. Exposures can be occupational, dietary, accidental, or intentional (e.g., suicide).

- Pesticides work via numerous different mechanisms in pest species as well as in humans and animals.
- The persistent organochlorine insecticides have been replaced by organophosphate compounds. These organophosphate insecticides are now being replaced by pesticides such as pyrethrins that are even of lower toxicity and are not very persistent.
- Industrial hygiene standards, such as OSHA PELs, and ACGIH TLVs and BEIs, exist for a number of pesticides.

REFERENCES AND SUGGESTED READING

American Conference of Governmental Industrial Hygienists (ACGIH). *Threshold Limit Values (TLVs) for Chemical Substances and Physical Agents and Biological Exposure Indices (BEIs)*. Cincinnati: ACGIH; 2012.

Baselt RC, Cravey RH. *Disposition of Toxic Drugs and Chemicals in Man*. 4th ed. Foster City: Chemical Toxicology Institute; 1995.

Berry C, La Vecchia C, Nicotera P. Review: paraquat and Parkinson's disease. Cell Death Differ 2010;17:1115–1125.

Blain PG. Organophosphorous poisoning (acute). Clin Evid 2011; 5:2102.

Bolt HM. Quantification of endogenous carcinogens. The ethylene oxide paradox. Biochem Pharmacol 1996;52:1–5.

Bond GG, Rossbacher R. A review of potential human carcinogenicity of the chlorophenoxy herbicides MCPA, MCPP, and 2,4-DP. Br J Ind Med 1993;50:340–348.

Buckley NA, Eddleston M, Li Y, Bevan M, Robertson J. Oximes for acute organophosphate pesticide poisoning. Cochrane Database Syst Rev 2011;2:CD005085. DOI: 10.1002/14651858.CD005085.pub2.

Coye MJ, Barnett PG, Midtling JE, Velasco AR, Romero P, Clements CL, Rose TG. Clinical confirmation of organophosphate poisoning by serial cholinesterase analyses. Arch Intern Med1987;147:438–442.

Daniell WS, Barnhart P, Demers LG, Costa DL, Eaton M, Miller, Rosenstock L. Neuropsychological performance among agricultural pesticide applicators. Environ Res 1992;59:217–228.

Delaplane KS. *Pesticide Usage in the United States: History, Benefits, Risks, and Trends*. Athens: Cooperative Extension Service, The University of Georgia College of Agriculture and Environmental Sciences; 1996.

Eddleston M, Buckley NA, Eyer P, Dawson AH. Management of acute organophosphate pesticide poisoning. Lancet 2008;371: 597–607.

Ellenhorn MJ. *Ellenhorn's Medical Toxicology: Diagnosis and Treatment of Human Poisonings*. 2nd ed. Baltimore: Williams and Wilkins; 1997.

Frankowski BL, Bocchini JA. The Council on School Health and Committee on Infectious Diseases. Clinical report—head lice. Am Acad Pediatr 2010. DOI: 10.1542/peds.2010-1308.

Gombe S, Ogada TA. Health of men on long term exposure to pyrethrins. East Afr Med J 1988;65:734–743.

Guzelian PS. Therapeutic approaches for chlordecone poisoning in humans. J Toxicol Environ Health 1981;8:757–766.

Guzelian PS. Comparative toxicology of chlordecone (Kepone) in humans and experimental animals. Annu Rev Pharmacol Toxicol 1982;22:89–113.

Hall SW, Baker BB. Intermediate syndrome from organophosphate poisoning. Abstract 103. Vet Hum Toxicol 1989;31:35.

Husain K, Ansari RA, Ferder L. Pharmacological agents in the prophylaxis/treatment of organophosphorous pesticide intoxication. Indian J Exp Biol 2010;48:642–650.

Ibrahim, MA, Bond GG, Burke TA, Cole P, Dost FN, Enterline PE, Gough M, Greenberg RS, Halperin WE, McConnell E, Munro IC, Swenberg JA, Zahm SH, Graham JD. Weight of the evidence on human carcinogenicity of 2,4-D. Environ Health Perspect 1991;96:213–222.

Krieger R. *Hayes' Handbook of Pesticide Toxicology*. Volumes 1 and 2. New York: Elsevier; 2010.

Larner AJ. Cholinesterase inhibitors: beyond Alzheimer's disease. Expert Rev Neurother 2010;10(11):1699–1705.

Lotti M. Cholinesterase inhibition: complexities in interpretation. Clin Chem 1995;41/42:1814–1818.

Lotti M, Moretto A, Zoppellari R, Dainese R, Rizzuto N, Barusco G. Inhibition of lymphocytic neuropathy target esterase predicts the development of organophosphate-induced delayed polyneuropathy. Arch Toxicol 1986;59:176–179.

Mattsson JL, Eisenbrandt DL. The improbable association between the herbicide 2,4-D and polyneuropathy. Biomed Environ Sci 1990;3:43–51.

Milby TH. Prevention and management of organophosphate poisoning. JAMA 1971;216:2131–2133.

MMWR. Fatalities resulting from sulfuryl fluoride exposure after home fumigation—Virginia. Morb Mortal Wkly Rep MMWR 1987;36:602–611.

Roberts JR, Reigart JR. *Recognition and Management of Pesticide Poisonings*. 6th ed. Washington: United States Environmental Protection Agency Office of Pesticide Programs; 2013.

Richardson RJ. Interactions of organophosphorous compounds with neurotoxic esterase. In: Chambers JE, Levi PE, editors. *Organophosphates: Chemistry, Fate and Effects*. New York: Academic Press; 1992. p 299–323.

Safe SH. Is there an association between exposure to environmental estrogens and breast cancer? Environ Health Perspect 1997;105 Suppl 3:675–678.

Sawada Y, Nagai Y, Ueyama M, Yamamato I. Probable toxicity of surface-active agent in commercial herbicide contained glyphosate. Lancet 1988;1:299.

Shore RE, Gardner MJ, Pannett B. Ethylene oxide: an assessment of the epidemiological evidence on carcinogenicity. Br J Ind Med 1993;50:971–997.

United States Environmental Protection Agency (USEPA). *Health and Nutrition Examination Survey II: Laboratory Findings of Pesticide Residues, National Survey*. Washington: USEPA; 1980.

United States Environmental Protection Agency (USEPA). *NHATS Broad Scan Analysis: Population Estimates from Fiscal Year 1982 Specimens*. Washington: Office of Toxic Substances; 1989.

United States Environmental Protection Agency (USEPA). *Worker Protection Standard*. 40 Code of Federal Regulations Parts 156 and 170. Washington: USEPA; 2014.

United States Environmental Protection Agency (USEPA). 1997. *Office of Pesticide Programs*, Human Health Benchmarks for Pesticides. Available at http://iaspub.epa.gov/apex/pesticides/f?p=HHBP:HOME. Accessed July 30, 2014..

United States Environmental Protection Agency (USEPA). Office of Pesticide Programs, Chemicals Evaluated for Carcinogenic Potential; 2013.

USEPA. 2012. Types of pesticides. Available at http://www.epa.gov/pesticides/about/types.htm. Accessed September 26, 2012.

USEPA. *Pesticides Industry Sales and Usage. 2006 and 2007 Market Estimates*. Washington: Biological and Economic Analysis Division. Office of Pesticide Programs; 2011.

USEPA. 2012. *Human Health Benchmarks for Pesticides*. EPA-882-F-12-001. Available at http://www.epa.gov/pesticides/hhbp/HHBP-fact-sheet.pdf. Accessed August 6, 2014.

Williams PL. Pentachlorophenol, an assessment of the occupational hazard. Am Ind Hyg Assoc J 1982;43(11):799–810.

Wirdefeldt K, Adami H-O, Cole P, Trichopoulos D, Mandel J. Epidemiology and etiology of Parkinson's disease: a review of the evidence. Eur J Epidemiol 2011;26:S1–S59.

Wood S, Rom WN, White GL, Logan DC. Pentachlorophenol poisoning. J Occup Med 1983;25:527–530.

World Health Organization. *Safety of Pyrethroids for Public Health Use*. Geneva: World Health Organization. Communicable Disease Control, Prevention and Eradication. WHO/CDS/WHOPES/GCDPP/2005.10; 2005.

18

PROPERTIES AND EFFECTS OF SOLVENTS AND SOLVENT-LIKE CHEMICALS

CHRISTOPHER M. TEAF, MICHELE M. GARBER, AND J. MICHAEL KUPERBERG

For purposes of this chapter, solvent classes and individual examples are presented, as well as selected substances (e.g., vinyl chloride, polycyclic aromatic hydrocarbons (PAHs)) not typically viewed as solvents but that are addressed in conjunction with solvents or are major elements of solvent mixtures. Information is provided concerning solvent chemistry, toxicology, and potential health effects, including:

- Exposure potential and possible health hazards associated with industrial and environmental aspects of organic solvents
- Chemical properties of selected classes of organic solvents
- Target organ systems of selected classes of organic solvents
- Relationships among solvent chemical structures and potential effects
- Toxicology of selected solvent examples, including some substances not traditionally considered to be solvents, though they are used as such
- Important substances present as components of commercial solvents

18.1 GENERAL CONSIDERATIONS

Solvents are defined as chemicals capable of dissolving and dispersing other substances. That fundamental property explains the economic value and potential health significance in residential, commercial, and environmental circumstances for products used as cleaners, degreasers, chemical intermediates, and chemical delivery vehicles. Organic solvents are carbon-based substances representing a very large, chemically diverse group of industrially, occupationally, and ecologically important products. The first historically useful organic solvents (e.g., ethanol, other alcohols, wood resins, turpentine) were derived from vegetable sources, and development of cosmetics containing organic solvent components dates to ancient Egypt. Inorganic solvents (e.g., water, ammonia, inorganic acids) are not addressed in this chapter, but comprise a chemically and economically important group as well.

There are thousands of organic solvents and related compounds, both pure compounds and commercial mixtures (e.g., mineral spirits, naphthas). The large numbers often result in generic statements about what solvents are or what they are not, as well as generalization about toxicity. However, there are marked differences among even similarly structured solvents. For example, low-molecular-weight members within a chemical class often exert greater toxicity than higher class members, due to differences in volatility, water solubility, and ability to cross biomembranes. That diversity demands care when comparing chemical and toxicological properties.

18.2 EXPOSURE POTENTIAL, MONITORING, AND RISK CONSIDERATIONS

Exposure Routes

Solvent exposure in the human environment is common, due to widespread use in occupational settings of many large industries and small commercial operations (e.g., automotive shops), and common potential for exposure via household

Principles of Toxicology: Environmental and Industrial Applications, Third Edition. Edited by Stephen M. Roberts, Robert C. James and Phillip L. Williams.

products. Such exposure can be by oral (e.g., drinking water), dermal (e.g., bathing, industrial direct contact), or inhalation routes (e.g., airborne workplace exposure, residential vapor intrusion), or a combination, depending upon the circumstances.

This chapter addresses airborne exposure potential in detail, due to the volatile nature of many solvents. However, as noted, solvent exposure is possible by multiple routes. For example, in situations where groundwater with solvent contamination is used for household purposes, there is the possibility of ingestion from drinking the water, dermal exposure from bathing activities, and inhalation related to volatilized solvents during clothes washing, showering, cooking, and other chores. Similarly, in industrial operations that employ solvents, joint exposure may occur due to splashes onto the skin, inhalation of vapors in work areas, and hand-to-mouth transfer of the chemicals during eating, smoking, or other common activities. Recognition of these multiple potential pathways has led to workplace requirements involving respiratory protection and hygiene practices limiting exposure possibilities. Differences in absorption, and other route-specific considerations for solvents, are further addressed in Section 18.3.

Not only is it important to address potential exposures to individual solvent agents but also possible interactive effects of multiple exposures, since these are the rule, rather than the exception. Assessment of the magnitude of exposure is often complex and may require detailed evaluation of inhalation and/or dermal contact, as well as estimates of exposure frequency and duration. Some common volatile organic chemical (VOC) solvents that may be encountered indoors and outdoors include benzene, toluene, xylenes, alcohols, trichloroethylene (trichloroethene, TCE), and formaldehyde. Exposure to organic solvents and their constituents can occur for individuals who live near industrial facilities that presently use or have used these solvents, as well as in the home. Many solvents are found in groundwater, soils, air, and other environmental media near National Priorities List (NPL) sites that are managed under Superfund (Comprehensive Environmental Response, Compensation, and Liability Act, aka CERCLA). Much of this is related to past storage, handling, and disposal practices at numerous industrial facilities. As of September 2014, there were over 1300 listed NPL sites and another 51 proposed sites in the United States, many of which are significant for the presence of variable solvent contamination profiles in soils and groundwater.

Industrial Exposures

There have been extensive advances in worker protection standards in the United States; however, industrial solvent exposures remain of global health interest for workers, many of whom are unfamiliar with potential hazards associated with exposures. The Occupational Safety and Health Administration (OSHA) has concluded that millions of U.S. workers are exposed to solvents in varying degrees on a daily basis. In some professions (e.g., painters), nearly all workers may have some degree of exposure, though education and protective measures (e.g., ventilation, spraybooths, respirators), coupled with introduction of water-based paints and adhesives, have reduced such exposure. Japan reports that smaller industries use nearly five times as much solvent volume compared to large enterprises, and they have experienced shifts in solvent use from aromatics to alcohols.

Exposures to organic solvents in an occupational setting also may occur in research laboratories, hospitals, and dry cleaners. In histology labs, formaldehyde (preservative) and xylenes (clearing agents) are the most common chemicals used and historically have been associated with pulmonary damage. Recently, housekeeping practices in the healthcare industry, which employ solvent-based cleaning products, have been identified as an occupational risk due to the asthma-like symptoms and dermatitis reported by some workers.

A limited epidemiologic database suggests that exposure to solvents and volatile chemicals in service station workers may be associated with restrictive airway disease, but demonstration of persistent lung complications has been inconsistent. A relationship also has been postulated for exposure to some chlorinated solvents during automotive degreasing work and an increase in risk of non-Hodgkin's lymphoma (NHL); however, that literature also is inconsistent. Ubiquity of solvents in industry, and potential for concurrent or sequential exposure to multiple chemicals, complicates the distinction between chemical causation of adverse effects and simple association of adverse effects with exposure in time and place. Relationships between simple association and chemical causation are discussed elsewhere in this book (see Chapter 21).

Industrial practices that result in the evaporative loss of volatile solvents (e.g., metal degreasing, application of surface coatings, chemical separations) are of particular interest in an exposure context. Protective equipment, engineering controls, and work practices can be effective in limiting exposures, but careless or inexperienced handling of solvents may still occur in small facilities (e.g., automobile service and repair, metal fabricators) and during activities in large and otherwise well-run factories and service industries. Methods used for worker exposure characterization and quantification are discussed elsewhere in this book (see Chapter 22). Solvent exposure potential varies among individuals, as well as over time for a specific individual, based on job type, workplace duties, and schedule. Within possible occupational exposure groups, it is important to consider sensitive individuals and populations when evaluating exposure and risk, including those with preexisting health conditions such as weakness of the immune system.

Occupational airborne guidelines that are designed to control exposures to solvents and other materials in the workplace may be expressed in units of volume/volume (e.g., parts per million (ppm)), as well as units of mass/volume (e.g., mg per cubic meter (mg/m^3)). For vapors and gases, these data if expressed in either form may be interconverted according to the following expressions:

$$X \text{ ppm} = \left[\frac{(Y \text{ mg/m}^3)}{\text{MW}} \right] (24.45)$$

where

X ppm (or ppb) = concentration in units of volume/volume
Y mg/m^3 (or µg/m^3) = concentration in units of mass/volume
MW = molecular weight of the chemical
24.45 = molar volume of an ideal gas at standard temperature and pressure

Rearranging this expression provides an opportunity to convert airborne concentrations that are expressed with different units in the other direction as well, as follows:

$$Y \text{ mg/m}^3 = \frac{(X \text{ ppm})(\text{MW})}{24.45}$$

For dose estimates, units of mg/m^3 are useful in conjunction with inhalation rates (units of m^3/h or m^3/day) to determine chemical intake in risk calculations. These unit conversion relationships do not apply for dusts, aerosols, and other nongaseous forms.

Household and Other Exposures

As noted, in addition to what are considered conventional industrial exposures, potential exposures to household products containing solvents and nonoccupational handling of petroleum products remain significant sources of exposure to hydrocarbon solvents. Residential exposure by all routes can occur when practicing hobbies (e.g., paints, thinners, adhesives), conducting home repairs (e.g., paints varnishes, thinners), using household cleaning products (alcohols, ethers, floor strippers, tub/tile cleaners), using fuels in lawn equipment (gasoline, diesel), and during recreational abuse of volatile inhalants, known as "huffing," in which a variety of solvents and pressurizing agents are used to attain intoxication-related euphoria, delusions, sedation, and hallucinations. Chronic inhalant abuse may lead to adverse neurologic conditions.

18.3 BASIC PRINCIPLES AND REVIEW OF TARGET ORGAN SYSTEMS

Structural variability and the range of physical/chemical properties exhibited by organic solvents limit the number of generalizations that can be made regarding physiological effects and exposure hazards. However, because of their common industrial, commercial, and household use, often in large quantities, it is useful to discuss some fundamental characteristics that are common to the principal classes of organic solvents. Table 18.1 summarizes selected important physical/chemical properties for solvents discussed in subsequent chapter sections. Of particular interest are vapor pressure (i.e., volatility) and water solubility, since these properties greatly influence environmental behavior and exposure potential. Many organic solvents are flammable/explosive, depending on their volatility, and have lower densities than water, except for some halogenated solvents. Solvent toxicity is greatly influenced by the number of carbon atoms, whether as a saturated or unsaturated molecule, as well as the chemical configuration (e.g., straight chain, branched, cyclic).

Table 18.2 presents occupational guidelines and standards for selected solvents and solvent constituents. These values include the health-based guidelines of the American Conference of Governmental Industrial Hygienists (ACGIH), termed threshold limit values (TLVs), as well as legally enforceable standards developed by the OSHA, termed permissible exposure limits (PELs). These guidelines and standards may be viewed as long-term protective levels, represented by a time-weighted average (TWA), as well as a protective value for a shorter time frame, termed a short-term exposure limit (STEL) or a ceiling (C) concentration. To the extent that they are available, carcinogen classifications from the United States Environmental Protection Agency (U.S. EPA), ACGIH, and National Toxicology Program (NTP) are included as well. Nearly half of the listed chemicals on Table 18.2 have no U.S. EPA carcinogenic classification, due to limitations to what the agency considers acceptable data.

In addition to the occupational air guidelines and standards, the ACGIH has published a series of biological exposure index (BEI) values for substances of specific industrial interest to guide the monitoring and control of health hazards. The BEIs are measures of a specific chemical or its metabolites in biological media (e.g., urine, blood, expired air). In principle, the BEI value is that which would be expected if airborne exposure regularly occurs at the TLV. The BEI is complementary to workplace air monitoring and can be a useful adjunct in situations where information is variable or contradictory. For example, in a situation where the BEI is exceeded but where workplace air levels are low, it would be prudent to investigate potential unrecognized peaks in air levels or to investigate the potential for unrecognized dermal exposure and absorption. Table 18.3 presents BEI values for selected solvents and related materials.

Table 18.4 provides definitions and differences among the most common occupational guidelines described in Table 18.2, as well as carcinogen classifications used by the U.S. EPA and ACGIH. As noted, the PEL is the legally

TABLE 18.1 Physical/Chemical Properties of Representative Solvents and Related Materials

Chemical	CAS #	Molecular Formula	Molecular Weight (g/mole)	Freezing or Melting Point (°F)	Boiling Point (°F)	Vapor Pressure (20 °C) (mm Hg)	Water Solubility (20 °C) mg/l	Specific Gravity (unitless)
Halogenated								
Carbon tetrachloride	56-23-5	CCl_4	153.8	−9	170	91	500	1.59
Chloroform	67-66-3	$CHCl_3$	119.4	−82	143	160	5,000 (25 °C)	1.48
Methylene chloride	75-09-2	CH_2Cl_2	84.9	−139	104	350	20,000	1.33
Tetrachloroethene	127-18-4	C_2Cl_4	165.8	−2	250	14	200	1.62
Trichloroethylene	79-01-6	C_2HCl_3	131.4	−99	189	58	1,000	1.46
Vinyl chloride	75-01-4	C_2H_3Cl	62.5	−256	7	2508	1,000 (25 °C)	2.21
Nonhalogenated								
Acetaldehyde	75-07-0	C_2H_4O	44.1	−190	69	740	Miscible	0.79
Acetone	67-64-1	C_3H_6O	58.1	−140	133	180	Miscible	0.79
Acrolein	107-02-8	C_3H_4O	56.1	−126	127	210	400,000	0.84
Aniline	62-53-3	C_6H_7N	93.1	21	363	0.6	40,000	1.02
Benzene	71-43-2	C_6H_6	78.1	42	176	75	700	0.88
Benzidine	92-87-5	$C_{12}H_{12}N_2$	184.3	239	752	low	400 (12 °C)	1.25
Carbon disulfide	75-15-0	CS_2	76.1	−169	116	297	3,000	1.26
N,N-Dimethylaniline	121-69-7	$C_8H_{11}N$	121.2	36	378	1	20,000	0.96
1,4-Dioxane	123-91-1	$C_4H_8O_2$	88.1	53	214	29	Miscible	1.03
Ethanol	64-17-5	C_2H_5OH	46.1	−173	173	44	Miscible	0.79
Ethyl acetate	141-78-6	$C_4H_8O_2$	88.1	−117	171	73	100,000 (25 °C)	0.9
Ethyl ether	60-29-7	$C_4H_{10}O$	74.1	−177	94	440	80,000	0.71
Ethylene glycol	107-21-1	$C_2H_6O_2$	62.1	9	388	0.06	Miscible	1.11
Formaldehyde	50-00-0	CH_2O	30.0	−134	−6	>760	Miscible	1.04
n-Hexane	110-54-3	C_6H_{14}	86.2	−219	156	124	20	0.66
Hydrazine	302-01-2	N_2H_4	32.1	36	236	10	Miscible	1.01
Isopropanol	67-63-0	C_3H_8O	60.1	−127	181	33	Miscible	0.79
Isopropyl ether	108-20-3	$C_6H_{14}O$	102.2	−76	154	119	2,000	0.73
Methanol	67-56-1	CH_4O	32.1	−144	147	96	Miscible	0.79
Methyl ethyl ketone	78-93-3	C_4H_8O	72.1	−123	175	78	280,000	0.81
Naphthalene	91-20-3	$C_{10}H_8$	128.2	176	424	0.08	30	1.15
Nitrobenzene	98-95-3	$C_6H_5NO_2$	123.1	42	411	0.3 (25 °C)	2,000	1.2
Nitromethane	75-52-5	CH_3NO_2	61.0	−20	214	28	100,000	1.14
Phenol	108-95-2	C_6H_6O	94.1	109	359	0.4	90,000 (25 °C)	1.06
Pyridine	110-86-1	C_5H_5N	79.1	−44	240	16	Miscible	0.98
Styrene	100-42-5	C_8H_8	104.2	−23	293	5	300	0.91
Tetrahydrofuran	109-99-9	C_4H_8O	72.1	−163	151	132	Miscible	0.89
Toluene	108-88-3	C_7H_8	92.1	−139	232	21	700 (23.3 °C)	0.87

TABLE 18.2 Occupational Exposure Limits for Selected Solvents and Related Materials

		ACGIH TLV (ppm)		OSHA PEL (ppm)		U.S. EPA Oral	Carcinogen Classification		
Compound	CAS #	TWA	STEL/Ceiling	TWA	STEL/Ceiling	RfD (mg/kg-day)	ACGIH	U.S. EPA	NTP
Acetaldehyde	75-07-0	NE	25	200	NE	NE	A2	B2	R
Acetone	67-64-1	250 (NIC)	500 (NIC)	1000	NE	0.9	A4	I	NE
Acrolein	107-02-8	NE	0.1	0.1	NE	0.0005	A4	I	NE
Aniline	62-53-3	2	NE	5	NE	NE	A3	B2	NE
Benzene	71-43-2	0.5	2.5	1	5	0.004	A1	A	K
Benzidine	92-87-5	NE	NE	NE	NE	0.003	A1	A	K
Carbon disulfide	75-15-0	1	NE	20	30	0.1	A4	NE	NE
Carbon tetrachloride	56-23-5	5	10	10	25	0.004	A2	L	R
Chloroform	67-66-3	10	NE	NE	50	0.01	A3	B2;L (acute); NL (low dose)	R
N,*N*-Dimethylaniline	121-69-7	5	10	5	NE	0.002	A4	NE	NE
1,4-Dioxane	123-91-1	20	NE	100	NE	0.03	A3	L	R
Ethanol	64-17-5	NE	1000	1000	NE	NE	A3	NE	K
Ethyl acetate	141-78-6	400	NE	400	NE	0.9	NE	NE	NE
Ethyl ether	60-29-7	400	500	400	NE	NE	NE	NE	NE
Ethylene glycol	107-21-1	10 (NIC)	100 (NIC)	NE	NE	2	A4	NE	NE
Formaldehyde	50-00-0	NE	0.3	0.75	2	0.2	A2	B1	K
n-Hexane	110-54-3	50	NE	500	NE	0.7	NE	II	NE
Hydrazine	302-01-2	0.01	NE	1	NE	NE	A3	B2	R
Isopropanol	67-63-0	200	400	400	NE	NE	A4	NE	NE
Isopropyl ether	108-20-3	250	310	500	NE	NE	NE	NE	NE
Methanol	67-56-1	200	250	200	NE	2	NE	NE	NE
Methylene chloride	75-09-2	50	NE	25	125	0.006	A3	L	R
Methyl ethyl ketone	78-93-3	200	300	200	NE	0.6	NE	I	NE
Naphthalene	91-20-3	10	NE	10	NE	0.02	A3	CBD	R
Nitrobenzene	98-95-3	1	NE	1	NE	0.002	A3	L	R
Nitromethane	75-52-5	20	NE	100	NE	NE	A3	NE	R
Phenol	108-95-2	5	NE	5	NE	NE	A4	I;D	NE
Pyridine	110-86-1	1	NE	5	NE	0.001	A3	NE	NE
Styrene	100-42-5	20	40	100	200	0.2	A4	NE	R
Tetrachloroethylene	127-18-4	25	100	100	200	0.006	A3	L	R
Tetrahydrofuran	109-99-9	50	100	200	NE	0.9	A3	S	NE
Toluene	108-88-3	20	NE	200	300	NE	A4	II	NE
Trichloroethylene	79-01-6	10	25	100	200	0.0005	A2	CaH	R
Vinyl chloride	75-01-4	1	NE	1	5	0.003	A1	A	K

Source: Compiled from ACGIH (2014).

K, known to be a human carcinogen; NE, not established; NIC, Notice of Intended Changes; PEL, permissible exposure limit; R, reasonably anticipated to be human carcinogens; RfD, reference dose; STEL/ceiling, short-term exposure limit or ceiling; TLV, threshold limit value; TWA, time-weighted average; see Table 18.4 for additional abbreviations.

TABLE 18.3 ACGIH Biological Exposure Index (BEI) for Selected Solvents and Related Materials.

Compound	CAS #	Recommended BEI	Sampling Medium	Sampling Time	ACGIH Notation
Acetone	67-64-1	50 mg/l	Urine	End of shift	Ns
Aniline	62-53-3	1 mg/l (tentative)	Urine	End of shift	Nq
		Nonquantitative	Released from hemoglobin	End of shift	Nq
		50 mg/l	p-Aminophenol[a] in urine	End of shift	Ns, Sq, B
Benzene	71-43-2	25 µg/g creatinine	S-Phenylmercapturic acid in urine	End of shift	B
		500 µg/g creatinine	*t,t*-Muconic acid in urine	End of shift	B
2-Butoxyethanol	111-76-2	200 mg/g creatinine	Butoxyacetic acid (BAA) in urine	End of shift	—
Carbon disulfide	75-15-0	0.5 mg/g creatinine	2-Thioxothiazolidine-4-carboxylic acid (TTCA) in urine	End of shift	Ns, B
Chlorobenzene	108-90-7	100 mg/g creatinine	Total 4-chlorocatechol in urine[b]	End of shift at end of workweek	Ns
		20 mg/g creatinine	Total p-chlorophenol in urine[b]	End of shift at end of workweek	Ns
2-Ethoxyethanol	110-80-5	100 mg/g creatinine	2-Ethoxyacetic acid in urine	End of shift at end of workweek	—
n-Hexane	110-54-3	0.4 mg/l	2,5-Hexanedione[c] in urine	End of shift at end of workweek	—
Methanol	67-56-1	15 mg/l	Urine	End of shift	B, Ns
2-Methoxyethanol	109-86-4	Nonquantitative	2-Methoxyacetic acid in urine	End of shift at end of workweek	Nq
Methylene chloride	75-09-2	0.3 mg/l	Urine	End of shift	Sq
Methyl ethyl ketone	78-93-3	2 mg/l	Urine	End of shift	Ns
Methyl n-butyl ketone	591-78-6	0.4 mg/l	2,5-Hexanedione[c] in urine	End of shift at end of workweek	—
Methyl isobutyl ketone	108-10-1	1 mg/l	Urine	End of shift	—
Nitrobenzene	98-95-3	5 mg/g creatinine	Total *p*-nitrophenol in urine	End of shift at end of workweek	Ns
		1.5% hemoglobin	Methemoglobin in blood	End of shift	B, Ns, Sq
Phenol	108-95-2	250 mg/g creatinine	Urine	End of shift	B, Ns
2-Propanol	67-63-0	40 mg/l	Acetone in urine	End of shift at end of workweek	Ns, B
Styrene	100-42-5	400 mg/g creatinine	Mandelic acid plus phenylglyoxylic acid in urine	End of shift	Ns
		0.2 mg/l	Venous blood	End of shift	Sq
Tetrachloroethylene	127-18-4	3 ppm	End-exhaled air	Prior to shift	—
		0.5 mg/l	Blood	Prior to shift	—
Toluene	108-88-3	0.02 mg/l	Blood	Prior to last shift of workweek	—
		0.03 mg/l	Urine	End of shift	—
		0.3 mg/g creatinine	o-Cresol in urine[a]	End of shift	B
1,1,1-Trichloroethane	71-55-6	40 ppm	End-exhaled air	Prior to last shift of workweek	—
		10 mg/l	Trichloroacetic acid in urine	End of workweek	Ns, Sq
		30 mg/l	Total trichloroethanol in urine	End of shift at end of workweek	Ns, Sq
		1 mg/l	Total trichloroethanol in blood	End of shift at end of workweek	Ns
Trichloroethylene	79-01-6	15 mg/l	Trichloroacetic acid in urine	End of shift at end of workweek	Ns
		0.5 mg/l	Trichloroethanol[d] in blood	End of shift at end of workweek	Ns
		Screening only	Blood	End of shift at end of workweek	Sq
		Screening only	End-exhaled air	End of shift at end of workweek	Sq
Xylenes	1330-20-7	1.5 g/g creatinine	Methylhippuric acids in urine	End of shift	—

Source: Compiled from ACGIH (2001) or updates as released.

B, background; Ns, nonspecific; Nq, nonquantitative; Sq, semiquantitative.

[a]With hydrolysis.

[b]After hydrolysis.

[c]Without hydrolysis; metabolite is specific to *n*-hexane and methyl *n*-butyl ketone.

[d]Without hydrolysis.

TABLE 18.4 Occupational Exposure Guideline Definitions

ACGIH—American Conference of Governmental Industrial Hygienists (ACGIH, 2011)	
TLV–TWA	Threshold limit value–time-weighted average. Time-weighted average concentration for a normal 8 h workday and a 40 h workweek, to which nearly all workers may be repeatedly exposed, day after day, without adverse effects
STEL	Short-term exposure limit. Defined as 15 min TWA exposure not to be exceeded during a workday. Concentration to which workers can be exposed continuously for a short period without suffering irritation, chronic/irreversible tissue damage, or narcosis to increase likelihood of injury, impair self-rescue, or materially reduce work efficiency
Categories for Carcinogenic Potential	
A1	Confirmed human carcinogen
A2	Suspected human carcinogen
A3	Animal carcinogen
A4	Not classifiable as a human carcinogen
A5	Not suspected as a human carcinogen
OSHA—Occupational Safety and Health Administration (ACGIH, 2011)	
PEL–TWA	Permissible exposure limit–time-weighted average. Concentration not to be exceeded during any 8 h work shift of a 40 h workweek
C	Ceiling concentration not to be exceeded any part of workday; if instantaneous monitoring not feasible, ceiling assessed as 15 min TWA
U.S. EPA—United States Environmental Protection Agency	
Cancer Classification (1986–1996)	
A	Known human carcinogen
B1	Probable human carcinogen. Based on human data
B2	Probable human carcinogen. Based on animal data
C	Possible human carcinogen
D	Not classifiable as to human carcinogenicity
Guidelines for Carcinogen Risk Assessment (Final in 2005)	
CaH	Carcinogenic to humans
L	Likely to be carcinogenic to humans
S	Suggestive evidence of carcinogenic potential
I or II	Inadequate information to assess carcinogenic
NL	Not likely to be carcinogenic to humans

enforceable standard for workplace air exposure and governs employers with regard to engineering practices, ventilation, and personnel protective equipment. It considers cost and technical feasibility in its development, as well as health considerations. The TLV is a guideline criterion that is based on health considerations only, derived from available animal and human toxicology information. In that respect, the TLV is analogous to the NIOSH recommended exposure limit (REL). TLVs and RELs often are used for occupational screening purposes. The U.S. EPA has established regulatory benchmark values for many of the substances discussed in this chapter [e.g., reference dose (RfD), cancer slope factor (CSF), and reference concentration (RfC) for air]. The **RfD** is defined as "An estimate of a daily oral exposure for a given duration to the human population (including susceptible subgroups) that is likely to be without an appreciable risk of adverse health effects over a lifetime. It is derived from a BMDL, a NOAEL, a LOAEL, or another suitable point of departure, with uncertainty/variability factors applied to reflect limitations of the data used." The oral **CSF** is defined as "An upper bound, approximating a 95% confidence limit, on the increased cancer risk from a lifetime oral exposure to an agent. This estimate, usually expressed in units of proportion (of a population) affected per mg/kg-day, is generally reserved for use in the low-dose region of the dose–response relationship, that is, for exposures corresponding to risks less than 1 in 100." The **RfC** is "an estimate (with uncertainty spanning perhaps an order of magnitude) of a daily inhalation exposure of the human population (including sensitive subgroups) that is likely to be without an appreciable risk of deleterious effects during a lifetime." Inhalation CSF values are available for some substances. These toxicological benchmarks often are used in calculating risk-based exposure targets and regional screening levels (RSLs) for solvents to be used in evaluation of risks posed by contaminated sites. They change periodically based on new information and can be acquired directly from online databases such as the U.S. EPA Integrated Risk Information System (IRIS).

As a point of comparison, Table 18.5 presents common ranges of concentrations encountered in environmental media and in occupational airborne circumstances.

TABLE 18.5 Concentration Ranges for Selected Solvents and Related Substances in Environmental and Occupational Circumstances

Chemical Class	Compound	CAS #	Common Occupational Exposure Range	Common Environmental Exposure Range: Air	Common Environmental Exposure Range: Water
Alkanes	*n*-Hexane	110-54-3	3–200 ppm	<50 ppb outdoor <10 ppb indoor	Low; BDL, 10 µg/l
Aromatic hydrocarbons—monocyclic	Benzene	71-43-2	0.05–83 ppb	0.02–34 ppb	Moderate; BDL, 100 µg/l
	Styrene	100-42-5	6.9–51 ppm	0.07–11.5 ppb indoor air 0.06–4.6 ppb outdoor air	Low; BDL, 20 µg/l
	Toluene	108-88-3	5–50 ppm	8 ppb indoor air	High; BDL, 1000 µg/l
	Xylenes	1330-20-7	0.5–1300 ppb	1–30 ppb outdoor air 1–10 ppb indoor air	High; BDL, 1000 µg/l
Aromatic hydrocarbons—polycyclic	Naphthalene		1.9–3.7 ppb	0.2 ppb or less outdoor air <1 ppb indoor air	Moderate; BDL, 200 µg/l
Aliphatics—halogenated	Methylene chloride	75-09-2	1–1000 ppm	<1–11 ppb outdoor	Moderate; BDL, 200 µg/l
	Tetrachloroethylene	127-18-4	16.9–48.4 ppm	<10 ppb outdoor air <15 ppb indoor air 74–428 ppb indoor air, newly dry-cleaned clothes	Moderate; BDL, 200 µg/l
	1,1,1-Trichloroethane	71-55-6	Up to 1300 ppm	0.1–1 ppb outdoor air 0.3–4.4 ppb indoor air	Low; BDL, 200 µg/l
	Trichloroethylene	79-01-6	50–100 ppm	0.08–12 ppb outdoor air <5 ppb indoor air	High; BDL, 1000 µg/l
Aromatics—halogenated	Chlorobenzene	108-90-7	Up to 4 ppm	<1 ppb outdoor air	Moderate; BDL, 100 µg/l
Alcohols	Phenol	108-95-2	0.03–3.2 ppm	0.5–44 ppb outdoor air	Low; BDL, 1–20 µg/l
Aldehydes, ketones	Acetone	67-64-1	1.5–166 ppm	7 ppb outdoor air 8 ppb indoor air	Moderate; BDL, 100 µg/l
	Methyl ethyl ketone	78-93-3	0.3–11 ppm	<10 ppb outdoor air	Low; BDL, 20 µg/l
Esters, ethers, epoxides	2-Butoxyethanol	111-76-2	0.1–169 ppm	<10 ppb outdoor air	Low; BDL, 10 µg/l
Amines—aromatic	Aniline	62-53-3	<2 ppm	1–5 ppb indoor air	Low; BDL, 10 µg/l
Nitro compounds	Nitrobenzene	98-95-3	1 ppm	0.01–5.7 ppb outdoor air	Low; BDL, 20 µg/l
Sulfur-containing compounds	Carbon disulfide	75-15-0	1–47 ppm	<10 ppb outdoor air	Low; BDL, 20 µg/l

Absorption, Distribution, Metabolism, and Excretion

As noted, solvent exposure may occur via oral, dermal, and inhalation routes, as well as a combination of these. Absorption can occur from direct liquid contact, and abraded or cut skin may enhance dermal absorption. While dermal penetration of solvents typically is negligible at low air concentrations, the ACGIH and OSHA note that this route for some substances may be significant at high air concentrations, hence an occupational "skin" designation, applicable in confined spaces or areas where respiratory protection (e.g., use of air-purifying or air-supplied respirators) limits potential for inhalation. For example, exposure to 2-butoxyethanol vapor may cause dermal uptake that exceeds inhalation. Solvent absorption by the lungs is dependent on several factors, including air concentration, as well as ventilation rate, depth of respiration, and pulmonary circulation, all of which are influenced by workload.

Once absorbed, solvents may be transported to other areas of the body, including organs where biotransformation may occur, resulting in formation of metabolites. Significant route-specific differences exist between uptake and the potential for adverse effects from solvents. Intake via the oral and dermal routes causes absorption into the venous circulation, which allows what is known as "first-pass" degradation or clearance by enzymatic process of the liver. Inhalation intake results in absorption via the alveoli into the arterial circulation, which distributes absorbed solvents to various locations prior to hepatic metabolism. For this reason, as well as the fact that airborne concentrations may be quite high in some cases, the inhalation route often is of greater toxicological concern on a strict numerical dose comparison.

Many volatile solvents can be eliminated either as the parent compound or in metabolized form in exhaled breath, varying with workload. This can be independent of exposure route and is the basis for sampling expired air as a measure of occupational exposure (see Table 18.3). Less water-soluble substances (e.g., chlorinated solvents) may penetrate more deeply into the lungs.

Given heterogeneity among solvents, numerous potential metabolic pathways exist, as described elsewhere in this book. The P450 enzyme system and the glutathione pathways often are involved, catalyzing oxidative reactions and conjugation to form water-soluble substances excretable in urine or bile. Several pathways may exist for the biotransformation of a specific solvent, depending on exposure route and concentration. Excretion of metabolites (e.g., s-phenylmercapturic acid from benzene, trichloroacetic acid from TCE, mandelic acid from styrene) forms the basis for biological monitoring programs in exposure characterization. Although the liver is the primary site of metabolism, other organs (e.g., kidney, lung) exhibit biotransformation capacity as well. Formation and accumulation of metabolites can occur if initial steps of biotransformation are present, but not later steps. For example, aldehydes may be metabolized readily in the liver, while the same aldehyde may accumulate in the lung and cause pulmonary damage due to limited aldehyde dehydrogenase. Beyond generally beneficial aspects of biotransformation (i.e., detoxification) and excretion, solvent metabolism occasionally may generate products that are *more* toxic than the parent compound. Such metabolic "activation," also termed "bioactivation," and resultant reactive intermediates (e.g., epoxides and radicals) are considered responsible for some toxic effects of solvents, especially with chronic exposure. This is illustrated by TCE, which exhibits variable metabolism in some rodents with dose, and that metabolism may produce novel toxicity elsewhere.

Metabolic enzymes may be increased in activity ("induced") by previous or concomitant chemical exposures, such as therapeutic drugs, foods, alcohol, cigarette smoke, and industrial exposures. Competitive interactions, which may affect enzyme activity for solvents in industrial scenarios, also may influence toxicity, such that exposure to multiple chemicals is not always worse than individual exposures (e.g., toluene inhibits benzene metabolism and toxicity). Saturation of normal metabolic degradation pathways may cause qualitative shifts in metabolism to other pathways. While a normal pathway may cause detoxification, saturation or overwhelming of the pathway may cause "shunting" to activation pathways (e.g., 1,1,1-trichloroethane, *n*-hexane, perchloroethylene (PERC), and 1,1-dichloroethene).

Partition coefficients (Kp) for air–blood, fat–blood, and brain–blood can be calculated and used to describe differential solvent behavior in various tissues. Such values are useful to understand the preferential uptake and storage of solvents and to explain greater sensitivity of some organs to adverse effects.

Acute versus Chronic Effects

As discussed subsequently in greater detail by chemical class, solvents can have both acute, short-term effects based upon one or a few high-level exposures (e.g., irritation, neurological effects) and chronic effects that may be expressed in delayed fashion following prolonged exposures at lower levels (e.g., kidney disease, liver injury, cancer). Acute effects typically result from the rapid overwhelming of existing detoxification and excretion mechanisms, and they may be transient/reversible. Chronic effects often are related to repetitive, unrepaired, low-level damage that is accumulative over lengthy periods of exposure, and these effects are less likely to be reversible. The intermediate case, where more frequent or regular high-level exposures may occur, is difficult to predict, given individual sensitivities and resilience.

For some solvents (e.g., aldehydes, ketones), the ability to detect odor or irritation at low air concentrations is a

useful characteristic that can act as an "early warning" tool against acute and chronic exposure potential. However, for most industrially and environmentally important solvents (e.g., benzene, chlorohydrocarbons), odor does not serve as a sufficient warning property, and significant exposures can occur before they are detectable by smell. Avoidance of such exposure circumstances is a foundation of occupational and environmental monitoring programs.

Central Nervous System Effects

A common physiological effect associated with short- or long-term, high-level exposure to some organic solvents is depression of central nervous system (CNS) activity, causing general anesthesia, decreased brain/spinal cord activity, and lowered sensitivity to external stimuli, with unconsciousness or death as the severest consequence. Many solvents are lipophilic ("fat loving"), a feature typically coupled with a low affinity for water (hydrophobic). Those compounds tend to accumulate in lipid-rich areas, including blood, brain, and depot (storage) fats. Accumulation in or direct damage to nerve cells can disrupt normal nerve excitability and adversely affect nerve impulse conduction. While organic solvents with few or no functional groups generally are lipophilic and exhibit some limited degree of CNS-depressant activity, this property generally increases with carbon chain length. Changes in toxicity are most evident when larger functional groups are added to small organic compounds, since the increase in molecular size generally decreases water solubility and increases lipophilicity. In practical terms, this applies only to chemicals up to carbon chain length of about six. As molecular size increases beyond this point for a functional class (e.g., amines, alcohols, ethers), vapor pressure decreases and exposure issues (e.g., inhalation) decrease dramatically.

Unsaturated chemicals (i.e., where hydrogens have been lost, forming double or triple bonds between carbon atoms) typically are more potent CNS-depressant chemicals than their saturated (i.e., single bond) counterparts. Similarly, CNS-depressant properties of organic solvents generally are enhanced by an increasing extent of halogenation (e.g., chlorine, bromine) and, to a lesser extent, by addition of alcohol (OH) groups. For example, while methane and ethane are gases, are not anesthetics, and act as simple asphyxiants at high concentrations, both of the corresponding alcohol analogs (i.e., methanol and ethanol) are liquids and are potent CNS depressants. Methylene chloride has anesthetic properties, but chloroform ($CHCL_3$) is more potent than methylene chloride, and carbon tetrachloride (CCL_4) is the most potent anesthetic of the group.

Toluene encephalopathy has been reported in high-level inhalant abuse, characterized by severe neurotoxicity. Evidence is limited that such a syndrome occurs at low-level exposures and there is considerable doubt as to whether "solvent-induced chronic encephalopathy" exists in reasonable exposure circumstances. That syndrome, described in the late 1970s, was explored more fully in the 1980s, though recent reviews do not conclude that chronic low-level solvent exposure typically results in CNS or peripheral nervous system (PNS) injury.

To illustrate the nonspecificity of solvent neurotoxic effects, and related problems that may be faced by a health specialist in trying to diagnose poorly characterized exposure situations, the following constellations of symptoms are described for a few common agents. It should be noted that, in contrast to these acute effects, effects of chronic exposure to these agents may differ dramatically, as discussed elsewhere in the chapter:

- *Benzene*: euphoria, excitement, headache, vertigo, dizziness, nausea, vomiting, irritability, narcosis, coma, and death
- *Carbon tetrachloride*: conjunctivitis, headache, dizziness, nausea, vomiting, abdominal cramps, nervousness, narcosis, coma, and death
- *Methanol*: euphoria, conjunctivitis, decreased vision, headache, dizziness, nausea, vomiting, abdominal cramps, sweating, weakness, delirium, coma, and death

PNS Effects

The literature is clear that select organic solvents (e.g., *n*-hexane, methyl *n*-butyl ketone, carbon disulfide) can cause "distal axonal peripheral neuropathy," also often termed "dying-back axonopathy" as defined by degradation of the axon or nerve cell body. It typically affects the feet and lower legs before the hands and may be partially reversible if identified sufficiently early in the process. The condition is associated with weakness, alteration or loss of sensation, impairment of reflexes, and eventual generalized muscle wasting. Occupational development of the condition has been reported in painters, automotive technicians, and workers in the shoe and furniture manufacturing sectors. Development of the condition is slow, though it may be accelerated in cases of inhalant abuse, and deterioration arguably may progress beyond the time at which exposure ceases, though the mechanism for such progression is not described. It is necessary to diagnose these disorders carefully while keeping in mind the possibility of other potential etiologies for the peripheral neurological observations (e.g., trauma, lead exposure, inflammation, diabetes, autoimmunity, heredity).

Renal and Hepatic Effects

Nephrotoxic effects have been associated with both acute and chronic exposures to halogenated hydrocarbons (e.g., chloroform, tetrachloroethylene). The primary target of

chloroform is the proximal kidney tubule where metabolism via renal cellular P450 produces reactive intermediates that cause damage followed by protein and glucose leakage to the urine, as well as increased blood urea nitrogen (BUN) levels. PERC is one of a group of chemicals that cause serious sex-specific and species-specific α2-microglobulin nephropathy in animals. However, humans are not at risk of this particular nephropathy because we do not synthesize the α2-microglobulin protein. As described in subsequent sections, organic solvents known to damage the liver include ethanol and chlorinated hydrocarbons. Enzymes (e.g., P450) and other factors that affect solvent transformation play a role in determining the extent of hepatotoxicity associated with these and related chemicals. Factors that activate P450 increase toxicity of many chlorinated solvents, and those that inhibit (decrease) P450 activity tend to reduce toxicity. Coexposure to some alcohols and chlorinated solvents has been associated with exacerbation of toxicity compared with that expected from either agent alone.

TABLE 18.6 Relative CNS-Depressant and Irritant Potency of Selected Organic Solvent Classes

CNS-depressant potential:

Most	Halogen-substituted compounds
↑	Ethers
	Esters
	Organic acids
	Alcohols
	Alkenes
Least	Alkanes

Membrane and tissue irritant potential:

Most	Amines
↑	Organic acids
	Aldehydes = Ketones
	Alcohols
Least	Alkanes

Irritation of Tissues and Membranes

Cell membranes are composed principally of a protein–lipid matrix, and organic solvents may act to dissolve that matrix or extract the fat (lipid) portion from the membrane, causing membrane and tissue irritation. This "defatting" process, when applied to skin, causes drying, irritation, and cell damage. Injury to the lungs and eyes may be caused by similar processes. As described previously, addition of some functional groups to organic molecules predictably influences toxicological properties. For example, amines, organic acids, alcohols, aldehydes, and ketones may cause cell membrane damage by precipitation and denaturation of proteins following sufficient exposures.

As is true for CNS depression, irritation by unsaturated compounds generally is stronger than for the corresponding saturated analogs. As molecular size increases, irritant properties typically decrease, and solvent defatting action of the hydrocarbon portion is altered. Table 18.6 presents relative potency of selected functional groups regarding CNS-depressant effects and irritation. The rankings become less applicable as the larger, more complex, multisubstituted compounds are considered.

A number of solvents are inhalation irritants, though irritation often is restricted to the upper airways due to the high degree of water solubility of the substances and the tendency to dissolve in moisture of mucous membranes.

Cancer

As with toxicological evaluations of other potential solvent-related adverse effects, the complex nature of industrial exposure scenarios complicates the objective evaluation of malignancy attributed to a specific solvent. Thus, many occupational studies end up considering solvent exposure as a general "risk factor" for neoplasia but stop short of establishing "cause and effect" relationships. While in some occupational exposure circumstances it is possible to establish a causative element of human cancer risk for industrial activities (asbestos and mesothelioma in some shipyard employees; angiosarcoma in concentrated vinyl chloride exposure), such is not the case for most solvents. There is historical documentation for benzene as a human carcinogen under well-described intense exposure circumstances. Multiple factors may be responsible for the observed effects, but metabolism of benzene to one or more reactive metabolites (e.g., epoxides) is likely to be responsible for the myelotoxicity. A complementary hypothesis suggests that a depressant effect by benzene or its metabolites on cell-mediated immunity may influence the carcinogenic process. It is interesting that the substituted benzene analogs toluene and xylenes are noncarcinogenic, while styrene (or vinylbenzene) forms reactive metabolites, as does benzene, notably styrene oxide. Styrene recently was listed as reasonably anticipated to be a human carcinogen in the NTP 12th Report on Carcinogens. That listing was based on limited evidence in human studies, sufficient evidence in animals, and supporting data on carcinogenesis mechanisms.

Some chlorinated solvents, such as carbon tetrachloride, chloroform, tetrachloroethylene (PERC), TCE (trichloroethylene), and vinyl chloride, exhibit carcinogenic potential, notably hepatic tumors, in animals. Carcinogenic potential associated with TCE exposure has been of interest since the mid-1970s, when the National Cancer Institute (NCI) reported increases in liver cancer in male mice receiving TCE by gastric intubation. TCE, like some other chlorinated hydrocarbons, exhibits limited, inconsistent,

and controversial mutagenic activity in bacterial test systems after microsomal activation, so the mutagenic effect is likely dependent on the products of metabolism, which has influenced recent interest in actual TCE potency. In the recent 2011 *Toxicological Review*, the U.S. EPA characterizes TCE as "carcinogenic to humans" based upon procedures outlined in the *U.S. EPA 2005 Guidelines for Carcinogen Risk Assessment*. That classification relies on evidence of a putative causal association between TCE exposure in humans and kidney cancer, in addition to carcinogenicity in the liver and lymphoid tissues. Other groups (e.g., ACGIH) continue to classify TCE as A2 ("suspected human carcinogen"), while the International Agency for Research on Cancer (IARC) considers TCE in class 2A "probably carcinogenic to humans." PERC is considered "likely to be a human carcinogen by all routes of exposure" as published in the 2012 *Toxicological Review* released by the U.S. EPA.

Apart from leukemogenic effects associated with chronic benzene exposure, the literature available to document specific cancer hazards from exposure to organic solvents is inconsistent, though some epidemiological observations have been published regarding cancer and chlorinated solvent exposure. For example, Hodgkin's lymphoma and NHL arguably have been linked to occupational exposure to some aliphatic, aromatic, and chlorinated solvents. Recently, a study of nearly 15,000 aircraft maintenance workers with exposure to TCE and other solvents reported a *decrease* in overall cancer mortality but a small excess in NHL, multiple myeloma, and bile duct cancer. The U.S. EPA inhalation unit risk and oral CSF for TCE were predicated partly on NHL risk.

Reproductive Effects

For most solvents, reproductive effects are not among the most sensitive measures of toxic effects. Reproductive toxicity has been reported following exposure to 2-methoxyethanol (2-ME), 2-ethoxyethanol (2EE), and *n*-hexane. 2-ME may damage the testes and cause infertility at extreme doses. It also has been shown to be teratogenic in animal studies, as has exposure to 2-bromopropane, carbon disulfide, and ethylene glycol monomethyl ether. Reproductive dysfunction was reported following exposure at high levels to toluene and other petrochemicals, though more sensitive measures of exposure occur at lower levels. Toluene abuse in pregnancy can lead to teratogenic effects labeled "fetal solvent syndrome." Organic solvent exposure in pregnant women reportedly has been associated with poor cognitive development and neuromotor functioning in childhood, as well as low childhood birth weight. Retail products containing ethanol require federal warning labels directed at notification of potential birth defects and other elements of reproductive consequences.

18.4 TOXIC PROPERTIES OF REPRESENTATIVE ALIPHATIC AND ALICYCLIC HYDROCARBON SOLVENTS

Aliphatic Solvents

Alkanes (C_nH_{2n+2}) The chemical class of the saturated aliphatic hydrocarbons (alkanes, also termed paraffins) has many members and generally ranks among the least potentially toxic solvents based on acute effects. The group is comprised of straight chain or branched hydrocarbons containing only single bonds. Vapors of these solvents are mildly irritating to mucous membranes at the high concentrations required to induce their relatively weak anesthetic properties. The four chemicals in this series with the lowest molecular weight (i.e., methane, ethane, propane, butane) are gases with negligible toxicity, and their hazardous nature is limited almost entirely to flammability, explosivity, and basic asphyxiant potential. The chemicals are found in natural gas and can be released into the environment from the exhaust of gasoline and diesel engines, from municipal waste incinerators, and from many other combustion sources.

Higher-molecular-weight class representatives are liquid at ambient temperature (i.e., 20–25 °C) and have some CNS-depressant, neurotoxic, and irritant properties, but this is primarily a concern of the lighter, more volatile compounds in the series (e.g., pentane, hexane, heptane, octane, nonane). Higher-viscosity liquid paraffins, beginning with the 10-carbon compound decane, are fat solvents and primary irritants capable of dermal irritation and dermatitis after repeated, prolonged, or intense contact. Many of these substances occur naturally in crude and refined petroleum products.

Symptoms of acute poisoning are similar to those previously described as generally present in solvent intoxication (e.g., nausea, vomiting, cough, pulmonary irritation, vertigo/dizziness, slow and shallow respiration, narcosis, coma, convulsions, and death), with severity of symptoms dependent on the magnitude and duration of exposure. Ingestion of large liquid quantities (i.e., exceeding several ounces, about 1–2 ml/kg body weight) may produce systemic toxicity. If less than 1–2 ml/kg is ingested, the therapeutic approach typically involves a cathartic used in conjunction with activated charcoal to limit GI absorption. In either situation, pulmonary aspiration of the solvent is the primary concern from a medical perspective. Low-viscosity hydrocarbons attract particular attention in this context because their low surface tension allows them to spread over a large surface area, with the potential to damage the lungs after exposure to relatively small quantities. These chemicals may sensitize the heart to epinephrine (adrenaline), but that feature is not often a practical issue given the narrow range separating cardiac sensitization from fatal narcosis.

Chronic exposure to some alkanes (e.g., hexane, heptane) may cause polyneuropathy in humans and animals, characterized by lowered nerve conduction velocity and a "dying-back" degenerative change in the distal neuron sheath. Symptoms may include muscle pain and spasms, weakness, and paresthesias, characterized by tingling or numbness in the extremities. Common metabolites have been implicated as the causative agents, with 2,5-hexanedione and 2,6-heptanedione as toxic degradation products of hexane and heptane, respectively. Since the metabolites are oxidative products, first to the alcohol and then to the diketone, there is concern that structurally similar alcohols and ketones may produce similar neuropathies compared to the parent aliphatic hydrocarbons. Alkanes generally are not considered carcinogenic.

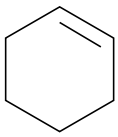
CONCENTRATION PROFILE FOR HEXANE IN HUMANS

Concentration (ppm)	Endpoint
50	TLV
130	Odor threshold
1000–2500	Drowsiness, fatigue
5000	Marked vertigo

Alkenes (C_nH_{2n}) Alkenes (or olefins), which are the double-bonded (i.e., unsaturated) structured analogs of the alkanes, exhibit qualitative toxicological similarities to the alkanes. The double bond typically enhances the irritant and CNS-depressant properties in comparison to alkanes, but this increase often is not sufficient to be of practical significance. For example, ethylene is a more potent anesthetic than its corresponding alkane (ethane) that acts as a simple asphyxiant. However, since a concentration greater than 50% ethylene is required to induce anesthesia, the potential for hypoxia and the explosive hazard are major drawbacks that preclude its clinical significance. Of greater toxicological interest is the observation that exposure to the unsaturated nature of the hexene and heptene series apparently largely *abolishes* the neurotoxic effects reported following chronic hexane or heptane exposure. This change may be related to substantive metabolic differences between the groups.

Alicyclic Solvents

Cyclopropane



Cyclohexene

Alicyclic hydrocarbons functionally are saturated alkane chains of which the ends have been joined to form a cyclic (ring) structure. They are distinct from aromatic rings (which are unsaturated), and toxicological properties resemble those of open chain analogs. They may exhibit anesthetic and CNS-depressant properties at high concentrations, but industrial experience indicates negligible chronic effects. Lower-molecular-weight alicyclics (e.g., cyclopropane) received some limited attention as surgical anesthetics, but the larger compounds (e.g., cyclohexane) are not as useful because the incremental difference between concentrations that cause narcosis and mortality is small. While there are qualitative similarities, the irritant qualities of cycloalkenes (cyclohexene and other cycloolefins) tend to be of greater concern than the unsaturated alicyclic analogs. Both ACGIH and OSHA have set 300 ppm for the TLV and PEL, respectively, for cyclohexene exposure.

18.5 TOXIC PROPERTIES OF REPRESENTATIVE AROMATIC HYDROCARBON SOLVENTS

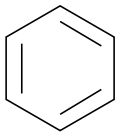
This class of solvents commonly referred to as "aromatics" is composed of one or more unsaturated six-carbon (phenyl) rings. The simplest member of the class, defined by lowest molecular weight, is the single-ringed analog termed benzene, followed by the aliphatic-substituted phenyl compounds (alkylbenzenes) and then the aryl- and alicyclic-substituted, multiringed benzenes. Diphenyl and polyphenyl compounds are represented in this class, which includes the polynuclear (polycyclic) aromatic hydrocarbons (PNAs, also termed PAHs), such as naphthalene, which are common as constituents of petroleum fuels, as well as other commercial products. Benzene and its alkyl relatives are important industrial compounds, with several billions of gallons of benzene annually produced or imported in the United States. Even larger quantities of several of the alkylbenzenes (e.g., toluene, xylenes) are produced. Benzene and the alkylbenzenes are common as manufacturing raw materials and solvents in the ink, dye, oil, paint, plastic, rubber, adhesive, chemical, drug, and petroleum industries. Most gasolines contain on the order of 1% benzene (i.e., 10,000 ppm), but the benzene content may range up to several percent. Some alkylbenzenes may be naturally present in or added to unleaded fuels to concentrations reaching 25–35% of the commercial product.

Aromatic hydrocarbons typically cause more tissue irritation than the corresponding molecular weight aliphatics and alicyclics. These phenyl compounds may cause primary dermatitis and defatting of the skin, resulting in tissue injury or chemical burns if dermal contact is repetitive or prolonged. Conjunctivitis and corneal burns have been reported when benzene or its alkyl derivatives are splashed as liquid into the eyes, and naphthalene has been reported to cause cataracts in animals at high dosages. If the aromatics are reaspirated into the lungs after ingestion (i.e., following

vomiting), they are capable of causing pulmonary edema, chemical pneumonitis, and hemorrhage. The inhalation of high concentrations can result in conditions ranging from bronchial irritation, cough, and hoarseness to pulmonary edema. Once absorbed into the systemic circulation, these hydrocarbons are demonstrably more toxic than aliphatics and alicyclics of comparable molecular weight. While CNS depression is a major acute effect of this class of compounds, its severe form differs fundamentally from that which is observed following exposure to the aliphatics. The aliphatic-induced anesthesia and coma is characterized by an inhibition of deep tendon reflexes. In comparison, aromatic-induced unconsciousness and coma are characterized by motor restlessness, tremors, and hyperactive reflexes, sometimes preceded by convulsions.

Representative members of the aromatic hydrocarbon family are profiled in the following section.

Monocyclic Aromatics



Benzene

Benzene is a colorless liquid with an odor generally described as pleasant or balsamic. Benzene is not synonymous with "benzine," and the latter term historically refers to a mixed-component, low-boiling-range, petroleum fraction composed primarily of aliphatic hydrocarbons. Due to its extensive historical industrial use, benzene has been studied perhaps more extensively than any other solvent. It can be toxic by all routes of administration at sufficient dosage; however, the acute inhalation LC_{50} in animals begins at about 10,000 ppm. This may be compared with observations in humans where lethal effects are observed at about 20,000 ppm within 5–10 minutes. Air concentrations near 300 ppm may produce vertigo, drowsiness, headache, nausea, and membrane irritation.

CONCENTRATION PROFILE FOR BENZENE IN HUMANS

Concentration (ppm)	Endpoint
0.5	TLV
1.5	Odor threshold
300–3,000	Dizziness, nausea, confusion
20,000	Acute lethality

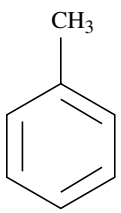
Ingested benzene exhibits comparatively greater systemic toxicity than corresponding aliphatic homologs, and the fatal adult human dose usually is reported to be on the order of 10 ml (125 mg/kg for a 70 kg person). Though CNS effects generally dominate over other systemic effects in short-term exposure circumstances, cardiac sensitization and cardiac arrhythmias also may be observed in severe intoxication. The pathology observed in acutely poisoned victims includes severe respiratory irritation, pulmonary edema and hemorrhage, renal congestion, and cerebral edema.

In pure liquid form, benzene is an irritant capable of causing dermal erythema, vesiculation (i.e., blistering), and dry, scaly dermatitis. Prolonged contact with benzene (or analogous alkylbenzenes) may result in lesions resembling first- to second-degree burns, and skin sensitization has been reported, as well as transient corneal injury.

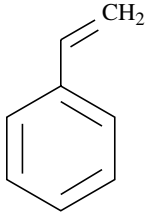
Benzene differs from most other organic solvents in that it is myelotoxic, with effects on blood-forming organs (e.g., marrow). Hematological effects from chronic exposure are variable including changes in red blood cell count, which may be 50% of normal, decreased hemoglobin, reduced platelets, and altered white cell counts, which often are reduced. In fact, in what later was recognized as misguided therapy, benzene was used in the early 1900s to decrease leukocyte counts in leukemia patients. Three stages have been identified in benzene-induced changes to blood-forming tissues. Initially, there may be reversible blood clotting defects and a decrease of all blood components, ranging from mild pancytopenia to aplastic anemia. The marrow may become hyperplastic and a stimulation of leukocyte formation may occur. While chronic benzene exposure may be best known for its putative link to specific leukemias, aplastic anemia actually is a more likely chronic observation. Several metabolites of benzene are implicated as the potential causative agents. Leukopenia and anemia in animals have been reported after chronic administration of hydroquinone and pyrocatechol, both benzene metabolites. However, benzene syndrome has *not* been observed in humans exposed to phenol, hydroquinone, or catechol.

Due to its presence in petroleum products, benzene is a common contaminant of groundwater at industrial and automotive service facilities. It is of interest not only due to its presence in drinking water but also as an air contaminant that may migrate from groundwater or soils to indoor air.

Benzene is classified as a known human carcinogen by the U.S. EPA, ACGIH, and others. Urinary muconic acid represents an accepted measure of industrial benzene exposure.



Toluene



Styrene

The aliphatic-substituted benzenes, also described by the term alkylbenzenes, include toluene (or methylbenzene), ethylbenzene, xylenes (or dimethylbenzenes), styrene (or vinylbenzene), cumene (or isopropylbenzene), and others. Toluene is a clear colorless highly flammable liquid with an odor similar to benzene. Styrene is a colorless to yellow oily liquid with a sweet, sharp odor. Unlike benzene, these substances are not generally considered to be carcinogens, and they do not generally cause effects in genotoxicity assays. However, toluene can exert a more powerful CNS-depressant effect than benzene, and human exposures at 200 ppm for periods of 8 hours generally will produce symptoms of fatigue, weakness, headache, and paresthesia. At 400 ppm, mental confusion is a symptom, and at 600 ppm, extreme acute fatigue, confusion, exhilaration, nausea, headache, and dizziness may result in a short time.

In addition to occupational exposure, concern recently has been directed toward the reports of intentional inhalant abuse of alkylbenzenes and alkylbenzene-containing products. Toluene is a primary solvent in paint thinners, spray paints, lacquers, and glues. Toluene inhalant abusers habitually expose themselves to levels exceeding 800 ppm, which causes euphoria but also leads to damage of the white matter in the brain. Acute toxicity of xylene isomers is qualitatively similar to toluene, though less potent. To a point, effects are reversible and become more pronounced as exposure time increases. Biological indicators of industrially important exposure have been developed for the alkylbenzenes, including urinary *o*-cresol (toluene), methylhippuric acid (xylenes), mandelic acid, and phenylglyoxylic acid (ethylbenzene, styrene).

Like benzene, toluene, xylenes and, to a lesser extent, styrene are components of refined petroleum products and are common contaminants of groundwater and soil at industrial and automotive service facilities. They are of interest not only due to their presence in drinking water, but also as air contaminants, since they may migrate from groundwater or soils to indoor air.

Polycyclic Aromatic Hydrocarbon (PAH) Compounds

This chemical class includes many members, all of which are cyclic-substituted benzenes. While this group is not technically in the solvent grouping, many of the PAHs are common components of petroleum fuels and some other solvent mixtures and are presented here for comparative purposes, given their toxicological significance.

PAHs are nonpolar, lipid-soluble compounds that may be absorbed via the skin, lungs, and digestive tract. Once absorbed, they may concentrate in organs with high fat content. They are metabolized by specific enzymes, which they also induce. These enzymes commonly are referred to as aryl hydrocarbon hydroxylase (AHH), or cytochrome P448. Since PAHs are composed of aromatic rings with limited available sites for metabolism, hydroxylation is the prevalent mechanism to initiate metabolism to more water-soluble forms that facilitate excretion. In this process, potentially toxic and carcinogenic epoxide metabolites may be formed. While the ubiquitous environmental presence of the PAHs suggests that regular exposure would more commonly lead to adverse effects, other routes of metabolism have been identified that appear to act as protective mechanisms by degrading reactive PAH metabolites. Similarly, natural or added constituents of foods such as flavonoids; selenium; vitamins A, C, and E; phenolic antioxidants; and food additives (e.g., butylated hydroxytoluene (BHT) and butylated hydroxyanisole (BHA)) all can exert protective effects. Recent evidence indicates that initial epoxide metabolites of PAHs are not the ultimate carcinogens because secondary metabolites of PAH are more potent mutagenic and carcinogenic agents, which also form DNA adducts more resistant to DNA repair. However, detailed discussion of these processes is provided elsewhere.

Naphthalene is the simplest PAH, containing two phenyl rings. It is a common fuel component, as well as a commercial moth repellent. Its white crystalline flakes have a strong odor. Naphthalene inhalation at sufficient concentration may cause headache, confusion, nausea, and perspiration. Severe exposures may cause optic neuritis and hematuria, with cataracts reported following naphthalene exposure in rabbits and at least one human case. Naphthalene is an irritant and rare hypersensitivity has occurred. The TLV, REL and PEL from the ACGIH NIOSH, and OSHA, respectively, are set at 10 ppm.

While PAHs can be acutely toxic, this characteristic generally is relevant only at doses sufficiently great that they are not of interest in an industrial or environmental setting. At high, acute doses, PAHs are toxic to many tissues including the thymus and spleen, while degenerative changes may ultimately be observed in the kidney and liver. For example, the noncarcinogen PAH acenaphthene, given in doses as high as 2000 mg/kg, produces only minor changes in the liver or the kidney and is relatively nontoxic when compared to the hematotoxicity produced by 100 mg/kg of dimethylbenzanthracene, a much more potent PAH. The teratogenic/embryotoxic effects of PAHs have only been documented for a few of the more potent, carcinogenic PAH compounds and then only in extreme exposure regimes in animal studies.

PAHs are commonly encountered in the environment, particularly in soils, water, and sediments, as they are common petroleum components and products of incomplete combustion of wood, wastes, fuels, and other substances. They also are created during cooking of meats and other food items. Several PAHs with greater than four rings (e.g., benzo(*a*)pyrene, benzo(*a*)anthracene, benzo(*b,k*)fluoranthene) are classified as possible carcinogens by environmental regulatory agencies. Occupational guidelines exist

for a category known as "coal tar pitch volatiles," which includes some of the PAHs. The ACGIH also has recommended BEI values for some of the PAHs.

18.6 TOXIC PROPERTIES OF HALOGENATED ALIPHATIC AND AROMATIC SOLVENTS

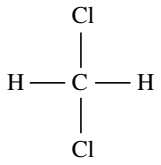
Many halogenated (e.g., chlorinated, brominated, fluorinated) solvents are widely used in industry for metal degreasing extraction processes, refrigerants, aerosol propellants, paint removers, and fumigants; as precursors in the manufacture of fluorocarbons; and as chemical intermediates in numerous chemical syntheses. They are largely nonflammable, and they generally exhibit potent anesthetic properties, several halogenated alkanes being the systemic anesthetic agents of choice in past or contemporary surgery (e.g., enflurane, halothane). Halogenated hydrocarbon anesthetics have negative effects on muscular rhythmicity and contractility as well as nerve conduction velocity (negatively chronotropic, inotropic, and dromotropic, respectively) at concentrations that typically are effective for anesthesia.

Halogenated compounds may exhibit strong dermal irritant effects. Brominated compounds are more toxic systemically and locally than chlorinated compounds, while fluorine replacement of the chlorine may decrease the observed toxicity. A drawback to the widespread use of halogenated alkanes is that some compounds (e.g., 1,2-dichloroethane) have been shown to induce liver cancer in rodent bioassays, though this has been demonstrated rarely or equivocally in human epidemiology studies. Many highly substituted halogenated alkanes are environmentally persistent.

Chronic exposure to some haloalkanes has been implicated in degenerative cardiac disease, resulting in investigation of cardiodepressant mechanisms, capacity for interference with energy production, and effect on intracellular calcium transport. Halogenated alkanes may sensitize the heart to endogenous epinephrine or to β-adrenergic drugs. Cardiotoxicity of low-molecular-weight halogenated hydrocarbons is considerably greater than that of low-molecular-weight unsubstituted analogs. Systemic toxicity to humans typically increases with increasing molecular size. The presence and degree of halogen substitution and the number of unsaturated bonds (e.g., progression from ethanes to ethenes) increase toxicity. Conversely, halogenated substitution of the aromatic ring may decrease systemic toxicity.

Halogenated Aliphatic Compounds

Halogenated aliphatics, particularly those with short alkyl chains and at least one chlorine or bromine atom, constitute a class of chemicals that are acutely nephrotoxic and hepatotoxic in animals. Evidence suggests that nephrotoxicity is related to metabolic products, rather than the parent haloalkane. Halogenated alkenes (e.g., TCE, PCE) are of historical and current occupational and regulatory interest.



Methylene chloride

Methylene chloride (dichloromethane (DCM)) is a volatile, colorless liquid with a sweet odor, has seen use as a solvent for many applications, including coating of photographic films, in aerosol formulations, and in paint processes where high volatility is desirable, and as a common "blowing" agent for synthetic foams. This volatility may result in high concentrations in poorly ventilated areas. Methylene chloride was used as a hair spray propellant, but this use was discontinued in 1989. It is a powerful solvent that is effective in dissolving cellulose esters, fats, oils, resins, and rubber.

DCM is the most water soluble and among the least toxic of the chlorinated methanes, with a predominant toxic effect of CNS depression, expressed as narcosis. Reported systemic effects following nonlethal exposures to methylene chloride include headache, giddiness, stupor, irritability, numbness, psychomotor disturbance, and increased blood carboxyhemoglobin. Prompt removal from exposure typically results in complete recovery. Methylene chloride is mildly irritating to the skin, and dermal absorption is not considered a significant threat to human health. Eye contact may be painful but is not likely to cause serious injury. Adaptation to methylene chloride vapors occurs with repeated contact, decreasing the ability to detect exposure.

CONCENTRATION PROFILE FOR METHYLENE CHLORIDE IN HUMANS

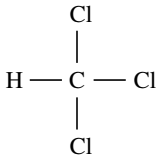
Concentration (ppm)	Endpoint
50	TLV
250	Odor threshold
2,300	Dizziness, nausea
10,000	Acute lethality

Historical industrial experience with methylene chloride has been remarkably free of serious adverse effects. Nephrotoxic and hepatotoxic potential is considered low in typical industrial uses. Reports of systemic injury are rare, though dermatitis has been reported due to common usage in paint remover formulations. Only a few deaths have occurred, all at extreme concentrations. DCM is metabolized to carbon monoxide, but symptoms of carbon monoxide poisoning, such as headaches, have not been a common feature of exposures, suggesting that carboxyhemoglobin levels alone are

not a good measure of the DCM toxic effects. Acute exposures may result in liver or kidney damage and cardiac sensitization in animal studies, though the latter is demonstrated only experimentally, limiting the practical significance of the findings.

Some studies with DCM in Ames mutagenicity tests show limited effects. Other genotoxicity tests are typically negative. NTP studies report clear evidence of carcinogenic potential at *high* doses of inhaled DCM in mice and female rats, though methylene chloride was not positive in a pulmonary tumor assay in mice. While regulated by the U.S. EPA and some occupational agencies as a potential carcinogen, DCM does not present a practical risk of carcinogenesis in humans at current acceptable exposure levels. The ACGIH recommends a BEI of 0.3 mg/l for urinary DCM.

DCM may be found in soils and groundwater at industrial or environmental contamination sites. Its volatility often limits its persistence. Complicating the interpretation of DCM environmental sampling results is the fact that DCM is commonly used in laboratories as an extraction solvent and may be found at low levels in blanks or other laboratory quality controls.



Chloroform

Chloroform (or trichloromethane) was used for many years as an anesthetic, solvent, insecticide, and chemical intermediate. It is a heavy, colorless, clear, nonflammable liquid with a characteristic pleasant odor. Due to delayed liver injury and reports of cardiac sensitization, its use as an anesthetic is obsolete. When there is a solid understanding of a chemical agent's mode of carcinogenic action, current *Guidelines for Carcinogen Risk Assessment* permit other approaches than the more historic, linearized default approach to estimating cancer risk. The U.S. EPA concluded that the dose–response of chloroform is nonlinear and is only likely to be carcinogenic under high-exposure conditions leading to cytotoxicity and hyperplasia in susceptible tissues. In the 2006 Stage 2 Disinfectants and Disinfection Byproducts Rule, the U.S. EPA calculated a maximum contaminant level goal (MCLG) for chloroform using the RfD. Sustained tissue toxicity, which is the key event in the cancer mode of action, is unlikely to occur at exposure levels below the RfD-based MCLG. Chloroform does not induce chromosome breakage or sister chromatid exchanges in human lymphocytes and failed to produce mutagenic changes in cultures of Chinese hamster lung fibroblast cells.

Chloroform often is produced at low levels during the chlorination/disinfection of water, particularly if raw water contains high levels of organic matter. Chloroform is one of the trihalomethanes (THMs) which are regulated in drinking water supplies.

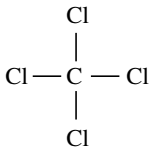
As with other volatile solvents, inhalation is the main exposure route for chloroform. Much available toxicological data results from interest in chloroform as a CNS depressant, and high concentrations may result in narcosis, anesthesia, and death. Rapidly absorbed and distributed, chloroform is exhaled unchanged or as carbon dioxide. It is the most cardiotoxic of the anesthetics, with effects that are the least reversible, and may cause permanent hepatic and renal injury. Hepatic necrosis was reported after ingestion of pharmaceutical preparations with 12.5% chloroform. Signs of severe exposure in humans include characteristic sweetish breath odor, dilated pupils, cold/clammy skin, excitation alternating with apathy, loss of sensation, loss of motor function, respiratory depression, cardiac sensitization, unconsciousness, coma, and death. Lethal chloroform effects may be delayed, resulting from associated kidney or liver damage. At exposure to concentrations below the anesthetic level, including occupational exposures, excitation passing into CNS depression may occur.

CONCENTRATION PROFILE FOR CHLOROFORM IN HUMANS

Concentration (ppm)	Endpoint
10	TLV
85	Odor threshold
20–200	Fatigue, nausea, vomiting
10,000	Clinical anesthesia

Chloroform is unusual among smaller chlorinated aliphatics in that it has been reported to be teratogenic and embryotoxic in animals, albeit at high exposure levels. It is found in umbilical cord blood at quantities greater than or equal to maternal levels.

Although toxicity of chloroform has been ascribed to solubility in cellular lipid, metabolism is necessary to explain the effects. Since chloroform does not form the free radical $\bullet CCl_3$, it would not be expected to be as toxic as carbon tetrachloride (see in the following text). However, mechanisms and paths of metabolism are not certain, and both enzymatic and nonenzymatic processes may be important. No metabolite has been identified in the blood or urine as a useful guide for evaluating occupational chloroform exposure.



Carbon tetrachloride

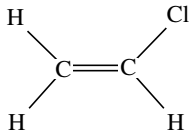
Carbon tetrachloride (or tetrachloromethane, carbon tet) saw historically widespread industrial uses that have declined since the 1970s, including fire extinguishers, refrigerants, metal degreasing, and semiconductor production and as a chemical raw material. Significant amounts were used as a grain fumigant. Carbon tet is an active insecticide and is effective in suppressing flammability of other flammable fumigants. The odor is one to which individuals may become adapted and is not a good warning of exposure. Carbon tet is not significantly teratogenic, but can be shown to be embryotoxic and fetotoxic to animals at very high maternal exposure levels. The substance historically has been regulated as a potential carcinogen by a number of occupational and environmental regulatory agencies, though its use is rare. Hepatic tumors were observed in multiple animal species (e.g., rat, mouse, hamster), while adrenal gland tumors were detected in mice. The 2010 U.S. EPA Toxicological Review for carbon tet describes it as "likely to be carcinogenic in humans," based on sufficient evidence in animals and inadequate evidence in humans.

Acute exposure to carbon tet may cause systemic effects including CNS depression, dizziness, loss of consciousness, dyspnea, cyanosis, proteinuria, optic vertigo, headache, mental confusion, incoordination, nausea, vomiting, abdominal pain, diarrhea, visual disturbances, ventricular fibrillation, kidney and/or liver injury, oliguria, albuminuria, edema, and anorexia. Clinical evidence suggests that carbon tet may cause ocular toxicity, although demonstration of retinal or optic nerve injury is lacking. Renal or hepatic injury may occur from a single acute exposure but is more likely following repeated exposures. The lower the exposure level, the greater the likelihood that injury will be focused in the liver. The CPY2E1 enzyme primarily is responsible for catalyzing bioactivation of carbon tet in humans, and concurrent intake of significant amounts of ethanol, which can increase the activity of CPY2E1, may greatly increase the probability of liver injury. In nonfatal poisoning, renal function occurs in three phases: (i) 1–3 days after exposure, oliguria stops, but creatinine and urea plasma concentrations remain elevated; (ii) there is a decline in these concentrations; and (iii) about 1 month after the initial injury, renal blood flow and glomerular filtration improve and renal function may recover after 3–6 months.

CONCENTRATION PROFILE FOR CARBON TETRACHLORIDE IN HUMANS

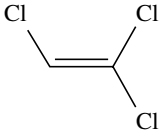
Concentration (ppm)	Endpoint
5	TLV
>10	Odor threshold
20–50	Headache, nausea, giddiness, possible liver injury

Application of pure carbon tet to human skin produces pain with erythema, hyperemia, and wheal formation followed by blistering that may result in secondary infection. Dermal absorption of carbon tet may present a potential problem based on limited human studies. In view of the potential hepatotoxicity of carbon tetrachloride, repetitive contact of the skin with liquid carbon tetrachloride should be avoided.

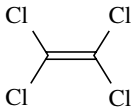


Vinyl chloride

Vinyl chloride (or chloroethylene) is not strictly a solvent, but is regularly encountered in environmental circumstances as a degradation product of TCE and PERC. It is a gas under ambient conditions and thus occupational exposures occur primarily via the inhalation route. It is a potent skin irritant with a high odor threshold (i.e., 3000 ppm). Contact with the liquid form, as it is usually handled under pressure and at low temperature, may cause frostbite, coupled with severe eye irritation. Vinyl chloride depresses the CNS, causing symptoms resembling mild alcohol intoxication. Lightheadedness, nausea, and dulling of visual and auditory responses may develop in acute exposures. Workers cleaning polyvinyl chloride reactor vessels have exhibited a triad of symptoms, including arthroosteolysis, Raynaud's phenomenon, and scleroderma. Chronic exposure may damage the liver and induce a highly specific liver cancer (i.e., angiosarcoma), which is an established risk for chronic exposures to vinyl chloride at the historical TLV of several hundred ppm. The current TLV set by the ACGIH and the PEL set by the OSHA are both 1 ppm. Experimental evidence in animals links vinyl chloride to tumor induction in a variety of organs, including the liver, lung, brain, and kidney, and to nonmalignant alterations (e.g., fibrosis, connective tissue deterioration). Vinyl chloride is classified as a known carcinogen by the EPA and ACGIH.



Trichloroethene



Perchloroethene

TCE (trichloroethene) and PERC (perchloroethene, tetrachloroethylene) historically have been among the most widely used of the chlorinated solvents. The toxicological literature is extremely large for these substances, and the reader is referred to the ATSDR Toxicological Profiles and the 2008 and 2011 Draft and Final U.S. EPA Toxicological Reviews for PERC and TCE, respectively, for a more detailed treatment. These substances exhibit a

generally low degree of acute toxic potential, and industrial use experience has been relatively good in cases where appropriate exposure controls were in place.

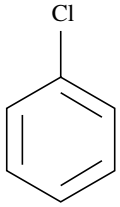
TCE has moderate water solubility and can therefore migrate through soils to the groundwater, where it is frequently detected at concentrations that exceed the low part per billion remediation target levels. TCE, as for most of the chlorinated solvents, is not readily degraded by common soil or groundwater microorganisms. The dominant fate of TCE in surface soils and surface water is volatilization.

CONCENTRATION PROFILE FOR TCE IN HUMANS

Concentration (ppm)	Endpoint
10	TLV
28	Odor threshold
100	Dizziness, nausea, confusion
10,000	Acute lethality

Although both TCE and PERC historically have been regulated as potential carcinogens by occupational and environmental regulatory agencies, ACGIH has reclassified TCE over the years. In 1993, ACGIH classified TCE as a group A5 (not suspected as a human carcinogen), and then in 2006 as a result of current epidemiology and case–control studies, it was changed to group A2 suspected human carcinogen, where it remains. In 1993, PERC was classified as a class A3 (animal carcinogen), where it remains today. Currently, the U.S. EPA has classified TCE as "carcinogenic in humans" and PERC as "likely to be carcinogenic in humans" in the final and draft versions, respectively, of each of their Toxicological Reviews. These documents have updated oral slope factors as well as unit risk factors listed. It is of interest to note that in some environmental conditions, both TCE and PERC may be degraded by sequential dechlorination to dichloroethenes and vinyl chloride, which was discussed previously.

Halogenated Aromatic Compounds



Chlorobenzene

Chlorobenzene is a colorless liquid with an almondlike odor that is manufactured as a solvent and used in production of other chemicals. Humans may be exposed by inhaling contaminated air, by ingestion of contaminated water/food, or by dermal contact with contaminated soils. Most occupational exposures occur from inhalation. Headaches and irritation of the upper respiratory tract or eye were reported following worker exposure to chlorobenzene in a glue preparation. Recent exposure to chlorobenzene can be quantified by measuring the total amount of 4-chlorocatechol and *p*-chlorophenol in the urine. Chlorobenzene currently is classified as an A3 animal carcinogen by the ACGIH, but its carcinogenic potential in humans is unknown.

Dichlorobenzenes

There are several important isomers of the dichlorobenzene molecule (e.g., *ortho*-dichlorobenzene, *para*-dichlorobenzene, *meta*-dichlorobenzene). 1,2-Dichlorobenzene (*o*-dichlorobenzene) is a pleasant smelling liquid used mainly as an intermediate in the synthesis of organic compounds and herbicides. 1,4-Dichlorobenzen (*p*-dichlorobenzene), the most important of the three, changes from a solid to a vapor when exposed to air, and commonly is used in deodorizer blocks. Exposure to these chemicals mostly occurs from breathing indoor air. In addition, they are used as industrial solvents, degreasing agents, insecticides, and fumigants. Dermal exposure can cause irritation followed by a burning sensation and skin redness. Eye and respiratory irritation occur in the workplace following exposures at about 100 ppm. There are only a few case reports of possible links between exposure to 1,2-dichlorobenzene and cancer, and it is not classifiable as a human carcinogen by any regulatory agency. *O*-, *m*-, and *p*-dichlorobenzenes are commonly detected at NPL sites identified by U.S. EPA, while OSHA and ACGIH regulate dichlorobenzene isomers in the workplace.

18.7 TOXIC PROPERTIES OF ALCOHOLS

In general, hydroxyl-substituted molecules (i.e., alcohols) are more powerful CNS depressants than their aliphatic analogs. In sequence of decreasing depressant potential, tertiary alcohols with multiple substituent -OH groups are more potent than secondary alcohols, which are more potent than primary alcohols. Alcohols also may exhibit irritant potential and generally are stronger irritants than similar organic structures that lack functional groups (e.g., alkanes) but are less irritating than corresponding amines, aldehydes, and ketones. Irritant properties of alcohols decrease with increasing molecular size. Conversely, potential for systemic toxicity *increases* with greater molecular weight, principally because water solubility is diminished and fat solubility is increased. Alcohols and glycols (or dialcohols) rarely represent serious hazards in the workplace, because practical vapor concentrations are usually less than irritant levels that prevent significant CNS effects

as well. In the environment, they typically are readily degraded and can act as a substrate that assists in the remedial degradation of other more recalcitrant compounds (e.g., chlorinated solvents).

Aliphatic Alcohols

$$H_3C-OH$$

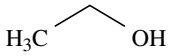
Methanol

Methanol (or methyl alcohol) is the simplest structural member of the alcohols (i.e., single carbon), occurring widely as an industrial solvent and raw material in manufacturing. It is a highly polar, flammable liquid with slight alcoholic odor that also is used as one of several intentional adulterants to "denature" ethanol, which in turn is used for cleaning, paint removal, and other applications. The denaturing process is designed to prevent ingestion by virtue of highly disagreeable taste. Methanol is of toxicological interest and industrial significance due to its unique toxicity to the eye, and it has received considerable attention from the medical community over the years due to misuse and accidental or intentional consumption. Methanol ingestion may have been responsible for 5–10% of all blindness in U.S. soldiers during World War II, presumably from unauthorized consumption. Methanol intoxication typically exhibits a combination of the following features:

- CNS depression, similar to or greater than that produced by ethanol
- Metabolic acidosis, caused by degradation of methanol to formic acid and other organic acids
- Ototoxicity, expressed as specific toxicity to retinal cells caused by formaldehyde, an oxidation product of methanol

Inebriation is not a prominent symptom of methanol intoxication, despite occasional consumption for that purpose. If ethanol is ingested in sufficient quantity, methanol poisoning may be considerably delayed and may even be averted completely. Consequently, the administration of ethanol after methanol intake makes it useful in treatment of methanol poisoning, based on competition for the alcohol dehydrogenase enzyme system. In order to increase oxidation of formate, a downstream product of methanol metabolism, folate therapy is used.

Acute methanol poisoning is characterized by headache, vertigo, vomiting, upper abdominal or back pain, dyspnea, restlessness, cold/clammy extremities, ocular disturbance, and diarrhea. Aside from well-described ocular effects, neurologic damage of various types may follow methanol poisoning. Death may be sudden, with accompanying inspiratory apnea and convulsions, or it may occur only after coma, depending on circumstances of exposure. The methanol oxidation rate is about 10–15% that of ethanol. Therefore, complete oxidation and excretion may require several days. An asymptomatic latent period of up to 24 hours may precede onset of adverse physiological effects. As little as 7 ml of methanol reportedly has caused blindness, and 70–100 ml may be acutely fatal. Occupational exposure values for methanol are set at 200 ppm for both the TLV and the PEL by the ACGIH and by OSHA, respectively. The ACGIH has also recommended a BEI of 15 mg/l for methanol measured in urine as an indicator of recent exposure.



Ethanol

Ethanol (or ethyl alcohol) is one of the largest volume organic chemicals in worldwide and U.S. production. It is used in many products including cleaners, ink, mouthwash, perfumes, pharmaceuticals, fuels, and beverages. Ethanol is a clear, colorless, flammable liquid with a sweet distinctive odor. Workplace exposure to ethanol typically occurs via inhalation, and illicit ingestion typically is not a problem due to the addition of what are known as "denaturants" to industrial ethanol that render it not fit for consumption due to taste and toxicity considerations. The denaturant typically is 5–10% methanol, but other denaturant additives include isopropanol, acetone, methyl ethyl ketone (MEK), methyl isobutyl ketone, fusel oil (butyl or amyl alcohol mixtures), and pyridine. In high concentration, ethanol acts as a mild to moderate local irritant, with the ability to injure cells by precipitation and dehydration. The CNS typically is affected most markedly. Initial apparent stimulation that accompanies ethanol ingestion results from altered activity in areas of the brain that have been freed of inhibition through the depression of control mechanisms. Ethanol increases the pain threshold considerably in most individuals even at moderate doses.

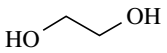
Dilation of cutaneous blood vessels, resulting in visible flushing, may accompany ethanol ingestion. Despite pervasive folklore, its use is not indicated during hypothermia from cold exposure. Another principal acute effect is cardiovascular depression of CNS origin. Ethanol may result in direct tissue damage at high chronic doses, producing damage to skeletal and cardiac muscles. Because ethanol increases gastric secretion at high doses, it has been linked to erosive gastritis, which can increase ulcer severity. It promotes fat accumulation in the liver in some circumstances, and chronic intake may lead to cirrhosis, liver cancer, and even death. Urine flow is promoted by inhibition of adrenal steroid and epinephrine release. Ethanol may exert a direct depressant action on bone marrow and may lead to a depression of leukocyte levels in inflamed areas, which may explain in part the

poor resistance to infection often reported in alcoholic individuals. Ethanol is classified as A3, confirmed animal carcinogen with unknown relevance to humans, by the ACGIH. The NTP classifies alcoholic beverages as known to be a human carcinogen given the presence of ethanol and other potential carcinogens (e.g., acetaldehyde, aflatoxins, urethane; see Table 18.2).

Aside from methanol and ethanol, high concentrations of propanols (or propyl alcohols), including isopropanol and *n*-propanol, may cause intoxication and CNS depression. They also have bactericidal properties. While both analogs have the same structural formulas (C_3H_7OH), isopropanol generally is less toxic than *n*-propanol, though both are more acutely toxic than ethanol. In humans, brief exposures to several hundred ppm of isopropanol in air generally cause mild irritation of the eyes, nose, and throat. *n*-Butanol (C_4H_9OH) potentially is more toxic than the lower-molecular-weight homologs, but it also is less volatile, which limits airborne exposure. Its symptoms may include eye, nose, and throat irritation; vertigo; headache; drowsiness; contact dermatitis; and corneal inflammation. No systemic effects from *n*-butanol typically occur at exposures less than 100 ppm on a regular basis. Allyl alcohol is more dangerous, and absorption through the skin can lead to deep pain. It may cause severe burns of the eye, and ocular symptoms include lacrimation, photophobia, and blurred vision. Allyl alcohol is metabolized by the liver to allyl aldehyde, a potent hepatotoxin. Alcohols may interact in industrial circumstances with chlorinated solvents via potentiation or synergism to enhance toxicity that would occur from either group alone.

Glycols

Larger alkyl-chain glycols, such as the double –OH group dihydroxy alcohols, exhibit a lesser acute oral toxicity compared to monohydroxy alcohols (single –OH group). They are rarely irritating to the eyes and skin and exhibit vapor pressures that are sufficiently low that toxic air concentrations are not usually observed at ambient temperature (e.g., 60–80 °F). Glycol solvents include ethylene glycol, diethylene glycol (DEG), and propylene glycol (PEG). Toxicological data on DEG are limited, but studies indicate renal failure as a finding following high doses. PEG is classified as "generally recognized as safe" by the U.S. Food and Drug Administration (FDA).



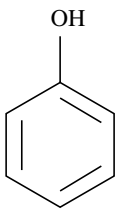
Ethylene Glycol

Ethylene glycol is a common member of the glycol family, used in detergents, paints, lacquers, pharmaceuticals, adhesives, and as a glycerin substitute. It is a colorless, clear, sweet, hygroscopic (i.e., attracts atmospheric moisture) liquid. More than 25% of ethylene glycol produced is used in coolant mixtures, most commonly in vehicle antifreeze applications. Common routes of exposure are dermal and accidental ingestion. A single oral dose on the order of 100 ml may be lethal in humans. Toxicity is principally due to metabolism of ethylene glycol, in the presence of calcium, to oxalic acid, which is toxic to the kidneys and may cause obstructive renal failure related to oxalate crystals. This occurs 24–72 hours after acute exposure. Humans exposed to ethylene glycol through inhalation do not exhibit renal damage. As with methanol, ethanol can be used as a competitive inhibitor of ethylene glycol by blocking aldehyde dehydrogenase-mediated metabolism. The sweet taste of ethylene glycol has resulted in many veterinary emergencies due to consumption of waste antifreeze.

Aromatic Alcohols

The aromatic alcohols (or phenols), in which the hydroxyl group is attached to a benzene ring, have the ability to denature and precipitate proteins in a manner similar to their aliphatic counterparts. This property makes phenol useful as a bacteriostatic agent at concentrations exceeding 0.2% and an effective bactericide in excess of 1.0%. However, it also renders these compounds corrosive, and severe burns may result from direct contact. Fatalities have resulted in individuals inadvertently splashed with liquid phenol due to its significant dermal absorption. Phenolic compounds also exhibit limited local anesthetic properties (hence their historical use in over-the-counter throat lozenges) and are CNS depressants at high concentrations.

Dihydroxy aromatics act much like the simple phenols, but their effects are largely limited to local irritation. The trihydroxy compounds may reduce the oxygen content of blood at sufficient exposure levels. Methylphenols (or cresols), while widely used in industrial applications, typically do not pose a significant inhalation hazard due to their relatively low vapor pressure and accompanying penetrating and objectionable odor. Their physiological effects are similar to those of phenol, and dermal exposure, if prolonged, may result in significant absorption, even to the extent that fatalities have been reported. Chlorinated phenols exhibit significant oral toxicity because of their direct inhibition of cellular respiration, and they are strong irritants. They also may produce muscle tremors, weakness, and, in overdose, convulsions, coma, and death.

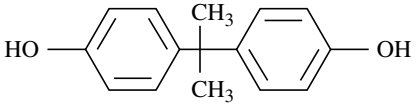


Phenol

Phenol, a colorless to pink hygroscopic solid with a sweet, tarry odor, can be cytotoxic to a variety of cells and tissues upon sufficient exposure, given its ability to complex with and denature proteins. Because it is easily absorbed and forms a loose complex with proteins, phenol may quickly penetrate the skin and underlying tissue, causing deep burns and tissue necrosis. This penetrating capacity, coupled with its nonspecific toxicity, renders it a serious handling hazard, and all routes of exposure should be controlled carefully. If splashed on skin, it produces redness and irritation with dermal injury ranging from eczema/discoloration and inflammation to frank necrosis or gangrene, depending on the degree and duration of exposure.

Following ingestion in concentrated solution, extensive local mucous membrane necrosis produced by phenol to the esophagus and stomach can cause severe pain, vomiting, and tissue damage, leading to shock and death. If inhaled at sufficient concentrations, phenol may cause chemical pneumonitis. Like many other solvents, phenol may cause tissue damage and necrosis in the liver and kidneys following absorption and systemic distribution. It is more potent in this respect than most organics, and phenol may produce degenerative, necrotic changes in the urinary tract and heart. A rapid fall in blood pressure may result from CNS-depressant properties of phenol on vasomotor control, as well as effects on the myocardium. In acute poisoning, death is usually related to respiratory failure. However, brief CNS stimulation and convulsions may occur. Lethal oral doses of phenol have been reported as low as 1–2 g. Phenol's dermal hazard is underscored by reports involving individuals on whom phenol accidentally was splashed. In one case, phenol spray over the thighs was fatal within 10 min of the exposure despite attempts to remove the clothing and to rinse the phenol off with water. The U.S. EPA concluded that data are inadequate to assess carcinogenic potential of phenol via the oral, dermal, or inhalation exposure routes. Urinary phenol represents an acceptable measure of industrial exposure.

Substituted phenolics include catechol (or *o*-dihydroxybenzene), resorcinol (or *m*-dihydroxybenzene), hydroquinone (or *p*-dihydroxybenzene), and cresols (or methylphenols), which have toxicological properties qualitatively similar to phenol. The alkyl substitutions tend to increase chronic toxicity, and all the substituted phenols are in some ways more chronically toxic than the parent phenol. Catechol may induce methemoglobinemia in addition to those toxic effects previously described for phenol.



Bisphenol A

Bisphenol A (BPA) is a high production volume (HPV) industrial chemical used since the 1960s to manufacture many hard plastics and metal-based food and beverage cans. It is a white to tan flake or powder with a mild odor. In 2008, the FDA conducted a review of the toxicity of BPA and concluded that food-related material containing BPA was safe. The NTP subsequently concluded that BPA may pose a risk to human development and reproduction, although at current exposure levels they only had "some concern" for effects on the brain, behavior, and prostate gland in fetuses, infants, and children, believed to have the highest estimated daily intakes of BPA. The NTP has "minimal concern" for workers exposed to high BPA levels in occupational settings. It has not undergone complete evaluation under the U.S. EPA IRIS program to determine human carcinogenic potential, nor have occupational exposure limits been set.

18.8 TOXIC PROPERTIES OF ALDEHYDES, KETONES, AND CARBOXYLIC ACIDS

Aldehydes

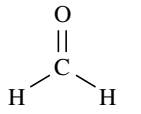
Aldehydes are primary irritants of the skin, eyes, and mucosa of the respiratory tract. These phenomena are most evident in lower-molecular-weight aldehydes and those with unsaturated aliphatic chains. Although a number of the aldehydes can produce narcosis, that effect rarely is observed because irritation that accompanies exposure generally serves as a sufficient warning property. Some of the aldehydes, such as fluoroacetaldehyde, can be converted to the corresponding fluorinated acids, giving them an extraordinarily high degree of systemic toxicity. The irritant properties of the dialdehydes have not been intensively studied. But, in some instances, concentrated solutions can be severe ocular and dermal irritants. The acetals and the aromatic aldehydes generally exhibit a greatly reduced potential for direct irritation.

An endpoint of toxicity that is common to aldehydes, but is not common to most other organic solvent constituents, is sensitization to the same or related substances. Formaldehyde is the most common agent among the aldehydes with respect to this problem, and sensitization reactions have been reported in persons who have been exposed to formaldehyde merely by wearing "permanent press" fabrics containing melamine-formaldehyde resins. Because their irritant effects limit inhalation exposure, the industrial use of aldehydes is relatively free of problems associated with other systemic organ toxicities of a serious nature. However, at sufficient concentration, respiratory damage is possible.

Unsaturated aldehydes (e.g., acrolein and ketene) are especially acutely toxic. The double bond in close association with the aldehyde functional group makes these compounds highly reactive with biological molecules and, therefore,

more toxic than the saturated analogs. For example, ketene and acrolein are on the order of 100 times more potent when measuring acute lethality by inhalation than either acetaldehyde or propionaldehyde. Reflecting this increase in potency is the fact that potential damage to the respiratory system is also more severe, resembling the deep lung damage of phosgene. Once absorbed, the systemic toxicity of the unsaturated aldehydes also can be more severe than that of the saturated members of the class.

Mutagenicity of some aldehydes (e.g., acrolein, formaldehyde, and acetaldehyde) suggests carcinogenic potential of these compounds. Some of these substances are regulated as possible carcinogens by various occupational and environmental agencies.



Formaldehyde

Formaldehyde is ubiquitous in the environment, as the simplest structural member of the aldehyde family. It is a colorless gas with a pungent irritating odor. It also is the most important aldehyde in commerce, with seven billion pounds produced in the United States annually, and is important from a regulatory perspective. Because of its high reactivity and instability in pure form, it generally is marketed in aqueous solution ranging from 37 to 50% formaldehyde ("formalin"). This product generally is diluted 1:10 for laboratory use, resulting in a usable concentration of approximately 4%. Formaldehyde also is commercially available in two other forms, a cyclic trimer, trioxymethylene, and a low-molecular-weight homopolymer, paraformaldehyde. The latter is used primarily in the plastic and resin industries, in synthesis of chemical intermediates, and, less frequently, in sealants, cosmetics, disinfectants, foot care creams, embalming fluids, corrosion inhibitors, film hardeners, wood preservatives, and biocides. Exposure to formaldehyde can occur in occupational settings as well as in schools, restaurants, office complexes, and mobile homes that may contain formaldehyde-based insulation, particleboard, wallpaper, plywood, carpet, and permanent press clothing.

Primary irritant effects of formaldehyde are considered the severest problem, and its local actions dominate the adverse effects observed following excessive exposure, in comparison to systemic effects that might otherwise occur. Table 18.7 describes the graded nature of the dose–response relationship for irritant and related effects from formaldehyde exposure. It also illustrates the range of effective concentrations reported for the human population. Formaldehyde can produce dermal sensitization in approximately 4% of the population, making it a common cause of occupational dermatitis. OSHA has set a PEL value of 0.75 ppm for occupational exposures. There is no present TLV.

Repeated dermal applications of high concentration solutions of formaldehyde induce sensitization in 5% of male subjects tested at low concentrations (≤2%). Formaldehyde is a common metabolite of normal cellular processes and serves as a cofactor in the synthesis of several essential biochemical substances. Tissue formaldehyde concentrations may reach several ppm in normal physiological conditions. The fatal oral dose of formaldehyde is estimated to be 60–90 ml of formalin (37%). Depending upon the dose, headaches, GI tract corrosion, pulmonary edema, fatty liver, renal tubular necrosis, unconsciousness, and vascular collapse may occur.

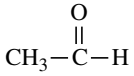
In the 2010 draft Toxicological Review of formaldehyde, a variety of interrelated noncancer health effects have been identified following airborne formaldehyde exposure, including irritation of the eyes, nose, and throat; upper respiratory pathology; reduced pulmonary function; asthma and atopy; neurologic and behavioral toxicity; reproductive and developmental toxicity; and immunotoxicity.

TABLE 18.7 Dose–Response Relationship for Formaldehyde Inhalation in Humans

Health Effects Reported	Formaldehyde Air Concentration
None	0–0.05 ppm
ATSDR minimal risk level (MRL)	0.008–0.04 ppm (chronic → acute)
NIOSH recommended exposure level (REL)	0.016 ppm
Odor detection	0.05–0.5 ppm
Eye sensitization and irritation	0.1–2.0 ppm
Upper airway irritation	0.10–2.0 ppm
ACGIH threshold limit value (TLV)	0.3 ppm (ceiling)
OSHA permissible exposure limit (PEL)	0.75 ppm (TWA)
Lower airway/pulmonary effects	5–30 ppm
Pulmonary inflammation/edema	50–100 ppm
Death	>100 ppm

Source: ACGIH (2001); ATSDR (1999); USEPA (2007); Cal OEHHA (1999).

While the exact mechanism of mutagenic actions is unknown, a recurrent finding in several test systems was that formaldehyde produces crosslinks within DNA that often are recognized and repaired by the DNA repair enzyme system. The mechanism for formaldehyde carcinogenicity appears to be a recurrent tissue injury and resultant hyperplasia caused by the high, irritating, and necrotizing exposures in test systems. Although some epidemiological studies have not found an association, other more recent epidemiological studies have found a positive link between formaldehyde exposure and nasopharyngeal cancer. In the 12th Report on Carcinogens, the NTP has changed the formaldehyde classification from the original 1981 *reasonably anticipated to be a human carcinogen* to classification of *known human carcinogen* based on human studies and data regarding mechanisms of carcinogenesis.



Acetaldehyde

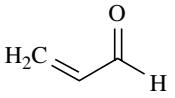
Acetaldehyde, a colorless, flammable liquid with a pleasant odor at dilute concentrations and a pungent/suffocating odor at high concentrations, is the next higher-molecular-weight aldehyde beyond formaldehyde and is also a common industrial chemical intermediate.

The primary source of occupational exposure has been fugitive emissions at fittings and pumps of manufacturing facilities. Acetaldehyde is a normal metabolite of mammalian ethanol metabolism and has been implicated by some authors as the source of the "hangover" associated with excessive ethanol consumption. It is less reactive than formaldehyde and, therefore, generally less irritating and toxic. It has been reported carcinogenic in animal tests, but its carcinogenic potential has not been investigated by extensive long-term oral studies. The U.S. EPA considers acetaldehyde a group B2 probable human carcinogen based on increased incidence of nasal tumors in rats and laryngeal tumors in hamsters following inhalation exposure, but no oral numerical estimate of risk has been derived. It is currently under review by the U.S. EPA. Acetaldehyde is embryotoxic and teratogenic in animal tests and is the presumptive key toxicant when ethanol has been tested in animal studies at sufficient dosage.

Acetaldehyde is rapidly absorbed and well metabolized in the liver. At 50 ppm in air, it is capable of eye irritation marked by conjunctivitis, and at 25 ppm, some people may feel irritation. It is detectable by odor, affording a considerable margin of safety, since the TLV ceiling based on health effects is 25 ppm. While it is a minimal industrial health hazard, it may represent an explosion hazard in some cases.

CONCENTRATION PROFILE FOR ACETALDEHYDE IN HUMANS

Concentration (ppm)	Endpoint
25 (ceiling)	TLV
0.21	Odor threshold
50	Eye irritation
100–200	Irritation of mucous membranes

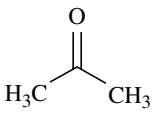


Acrolein

Acrolein is the unsaturated analog of propionaldehyde and is more toxic in comparison. It is a highly reactive, flammable liquid with a "choking," disagreeable odor. High vapor pressure and water solubility cause acrolein to be mobile in the environment, although degradation typically limits transport. Occupational exposure may occur through inhalation and dermal contact where acrolein is used or produced. It is toxic by all routes of administration and is capable of severe eye and pulmonary irritation. It primarily is retained in the upper respiratory tract following inhalation exposure. Since contact with the skin and eyes may produce necrosis, direct contact must be avoided. Given its acute toxicity, few data are available to evaluate chronic toxicological effects. The occupational guideline for acrolein is 0.1 ppm (TLV, ceiling, and PEL), not far below the level considered to be moderately irritating (0.25 ppm). No adequate human studies to determine carcinogenic potential of acrolein were found.

Ketones

The limited literature reports regarding serious injury suggest that most ketones do not present serious health hazards under typical circumstances, in part because they have fairly effective warning properties (e.g., identifiable odor). Ketones are recognized CNS depressants, but vapor concentrations sufficient to cause sedation also are irritating to the eyes and respiratory passages and are avoidable. Lower concentrations are easily inhaled and may reach levels that impair judgment. Toxic properties of ketones increase with increasing molecular weight, and unsaturated analogs are more toxic than saturated members. Industrially important exposures to ketones may be evaluated by relevant urinary measures (e.g., acetone, MEK, methyl isobutyl ketone).



Acetone

Acetone is a common industrial solvent and raw material, present in commercial and household products. It is a clear, colorless liquid with an aromatic odor. Daily ingestion of acetone for several days in doses as high as 15–20 g produced limited adverse effects (e.g., drowsiness). Skin irritation occurs only after repeated prolonged contact. Persons unaccustomed to acetone may experience eye irritation at 500 ppm and the OSHA PEL is 1000 ppm. While workers accustomed to daily exposure can easily tolerate high air concentrations, at 9,000–10,000 ppm, unambiguous irritation of the throat and lungs occurs. Studies of exposed employees with average exposure to acetone concentrations of 2000 ppm typically reveal no serious injury. Inhalation of high doses of acetone can lead to dryness of mouth and throat, dizziness, nausea, uncoordinated movement, speech impairment, and drowsiness. Animal studies have not demonstrated carcinogenic potential, and acetone is not classified as a carcinogen. Industrial exposures can be determined by measuring acetone levels in the urine.

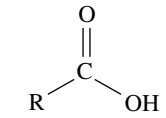
CONCENTRATION PROFILE FOR ACETONE IN HUMANS

Concentration (ppm)	Endpoint
250 (Notice of intended 2014 change)	TLV
13	Odor threshold
500	Eye irritation
>12,000	Unconsciousness, dizziness

MEK (2-butanone), a colorless, flammable liquid with acetone-like odor, is a common industrial raw material and solvent. Uses include production of 1,3-butanedione and MEK peroxide. It is detectable by odor at a few ppm in air, which represents a warning property at several hundred times less than its occupational exposure limit. High concentrations are irritating to the eyes, nose, and throat, and dermal or ocular irritation accompanies exposure to liquid splashed on the skin and in the eyes. CNS depression may result from prolonged exposure. The toxicological literature concerning MEK is extensive in the form of animal studies and human data, but the U.S. EPA determined that data are inadequate for an assessment of human carcinogenic potential. The TLV and PEL are both set at 200 ppm. The ACGIH also has set a BEI of 2 mg/l MEK in urine.

Acetophenone (phenylethylketone) was used historically as an anesthetic but is used today as a component of perfumes due to its persistent odor described as similar to orange blossoms or jasmine. It is a strong skin irritant, but not a potent CNS depressant. Eye contact to high concentrations may cause irritation and transient corneal burns. Acetophenone is regulated by the ACGIH for occupational exposure, and the U.S. EPA has determined that it is not classifiable as to its human carcinogenic potential.

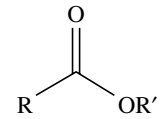
Methyl *n*-butyl ketone (or 2-hexanone) is a potent neurotoxin under some exposure circumstances and is metabolized to 2,5-hexanedione, the neurotoxic metabolite also attributed to hexane. Therefore, it may induce a polyneuropathy like that described previously for hexane. Quantification of 2,5-hexanedione in the urine can be used as a biological indicator of industrially important exposure. The TLV and PEL set by the ACGIH and OSHA are 5 and 100 ppm, respectively.



Carboxylic acid basic structure

As with aldehydes and ketones, the irritant properties of these compounds dominate observed effects and may mask CNS-depressant potential. Acidity (i.e., low pH) and irritancy decrease with increasing molecular size. Halogenation of carboxylic acids, particularly at the α-carbon, enhances the strength of the acid and makes a stronger irritant. Dicarboxylic acids and unsaturated carboxylic acids are comparatively more corrosive. For example, acetic acid (CH_3COOH) is only moderately irritating, but the unsaturated acrylic acid ($CH_2=CHCOOH$) and crotonic acid ($CH_3CH=CHCOOH$), as well as trichloroacetic acid (CCl_3COOH, a mammalian metabolite of TCE), may produce burns and tissue damage.

18.9 TOXIC PROPERTIES OF REPRESENTATIVE ESTERS, ETHERS, AND EPOXIDES



Ester basic structure

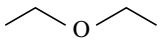
Esters often exhibit a fruity odor and are used as industrial solvents in making cellulose, fats, paints, and artificial perfumes. Esters of similar structure are more potent anesthetics than equivalent alcohols, aldehydes, and ketones but are weaker than ethers or halocarbons. Most esters are of limited skin and eye irritant interest, but low-molecular-weight esters are more potent irritants than alcohols and can cause eye irritation and lacrimation. Halogen substitution increases irritancy, and double bonds in the side chain may increase toxicity. Some unsaturated esters cause CNS *stimulation*, rather than CNS *depression*. Additional nonhalogen functional groups and double bonds between carbons in the

alkyl chain reduce vapor pressure and toxicity. Esters are degraded in the bloodstream by esterases to similar carboxylic acids and alcohols.

Phthalate esters, used as plasticizers, may produce CNS damage at very high concentrations, may be irritating, and have the capacity to act as convulsants or depressants. Due to the variety of ester structures, and dramatic effect that functional groups exert on toxicity, exposure to each compound requires specific review, and few generalizations are possible. Di(2-ethylhexyl) phthalate (DEHP, BEHP), a light-colored, viscous liquid with a slight odor, is classified as a potential carcinogen by some regulatory agencies, based on liver cancer observed in rodents. The U.S. EPA classifies it as in group B2 and the ACGIH assigns it an A3 animal carcinogen classification. The likely mechanism of action (repetitive tissue injury) renders this classification of limited human interest. ACGIH and OSHA set the TLV and PEL at 5 mg/m^3.

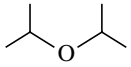
Ethers

Ethers have broad industrial and commercial usage. They are used in production of rubber, plastics, cosmetics, paints and coatings, refrigeration, and foods. They also were widely used in medicine. Ethers as a class are effective anesthetics, a property that increases with molecular size. The utility of this property is limited by irritant effects of ethers and the fact that they are easily oxidized or photodegraded to peroxides, which may be explosive and shock sensitive. Generally speaking, as ether chain length increases, oral and inhalation toxic potential decreases. In contrast, dermal penetrability and skin irritation tend to increase with increasing chain length.



Ethyl ether

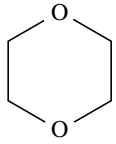
Diethyl ether (or ethyl ether), a colorless liquid with a characteristic sweet odor, is absorbed and excreted rapidly through the lungs. It is a potent anesthetic and was widely used in medicine at one time, but it is slightly irritating to the skin and contact with the eyes should be avoided. It produces anesthesia in humans at concentrations of 3.6–6.5% in air, but respiratory arrest occurs at 7–10%, providing a small margin of safety. Ethyl ether produces profound muscular relaxation by means of corticospinal and neuromuscular blockade; however, nausea and vomiting are common subsequent side effects, limiting its practical utility. Occupational exposure limits for diethyl ether are 400 ppm for both the TLV and PEL as set by the ACGIH and OSHA, respectively.



Isopropyl ether

Isopropyl ether (or diisopropyl ether) is comparatively more toxic than ethyl ether and causes irritation at much lower concentrations than those required to produce anesthetic effects, which limited its practical use as an anesthetic. It has a sharp, sweet, ether-like odor. In man, the odor is noticeably unpleasant at 300 ppm for 15 min, 500 ppm typically causes no irritation, and at 800 ppm, disagreeable irritation of the eyes and nose is noticeable. Consequently, the ACGIH and OSHA have set occupational exposure levels at 250 and 500 ppm, respectively.

Unsaturated ethers are more toxic than saturated analogs, produce anesthesia faster, and may cause liver damage (e.g., divinyl ether is more potent than ethyl ether). Halogenated ethers can cause severe irritation to the skin, eyes, and lungs. For example, chloromethyl ether vapors are painful at about 100 ppm. Chlorinated ethers also may be potent alkylating agents, and compounds such as bis(2-chloromethyl)ether are classified as carcinogens (see also Chapter 15). Aromatic ethers, in contrast, generally are less volatile, less irritating, and less toxic than the alkyl ethers.

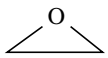


1, 4-Dioxane

1,4-Dioxane (or diethylene dioxide), a glycol ether, is used in a wide range of lacquers, paints, dyes, cosmetics, deodorants, stains, detergent products, and as a historical chemical stabilizing agent for some chlorinated solvents. In that application, it has been detected in circumstances related to solvent groundwater contamination, due to its very high water solubility, which can cause rapid migration. 1,4-Dioxane is distinguished from the wholly unrelated "dioxins" (e.g., 2,3,7,8-tetrachlorodibenzodioxin) in any discussions of toxicity, carcinogenesis, and environmental persistence. Liquid 1,4-dioxane is a painful irritant to the eyes and skin and can be absorbed dermally in significant quantities at high concentrations. Renal and hepatic damage may occur following prolonged exposure. While 1,4-dioxane is considered "likely to be carcinogenic to humans" based on animal data via oral exposure, human studies are inconclusive. Data regarding exposure to 1,4-dioxane by inhalation is being evaluated by U.S. EPA, and an assessment supplemental to a 2010 Toxicological Review is expected. The OSHA PEL is 100 ppm, while the ACGIH TLV is 20 ppm.

Soil, surface water, and groundwater contamination with 1,4-dioxane have been reported at areas near hazardous waste sites and industrial facilities. Its high solubility and volatility affect the removal process for soil and surface water contamination. 1,4-Dioxane may persist longer in the

groundwater, and if air stripping cannot be achieved, other remedial methods such as *ex situ* chemical oxidation, microbial bioreactors, and UV oxidation have been successful.



Epoxide basic structure

Epoxides are a subclass of ethers, meaning they contain the C–O–C configuration in a heterocyclic formation. They are considered reactive and easily polymerized. While some are useful as end products, most are used as intermediates and stabilizers. Ethylene oxide, the simplest aliphatic epoxide, is a colorless gas that is condensible to a clear liquid at 10 °C. Most human exposure occurs via inhalation associated with sterilization operations in hospitals and with fumigant use. The ACGIH TLV and OSHA PEL are set at 1 ppm. Epichlorohydrin is a volatile liquid used in the manufacture of other industrial chemicals. The ACGIH TLV is 0.5 ppm and it is classified as an A3 confirmed animal carcinogen with unknown relevance to humans based on animal studies. The OSHA set a PEL at 5 ppm. There is a skin notation warning of the danger of possible cutaneous absorption through the human skin.

18.10 TOXIC PROPERTIES OF REPRESENTATIVE NITROGEN-CONTAINING SOLVENTS

$R-NH_2$

Aliphatic amine

Nitrogen-substituted compounds have been used in a wide range of industrial applications, and aliphatic amines are among the most toxic organic chemicals. Both β-chloroethylamine and trichloronitromethane are strong irritants and have been used as effective chemical warfare agents. Most members are potent irritants and sensitizers, with potency related to the base strength of the amine group. Corrosive properties generally are unrelated to molecular size, in contrast to many other organic chemical groups (e.g., esters, ethers, alcohols) that exhibit decreasing irritancy as molecular size increases. The degree of substitution (i.e., primary, secondary, tertiary) has little effect on corrosiveness of the amine group. Thus, while the irritant nature of the other functional groups (e.g., alcohols, ethers, carboxylic acids) is decreased as the size of the organic portion of the chemical increases, the irritation of the amines is not affected.

One characteristic of amines influencing toxicity and handling hazards is efficient absorption by all routes. In addition to physical skin damage, they may cause systemic toxicity including methemoglobinemia, pulmonary hemorrhage, hepatic necrosis, nephrotoxicity, and cardiac degeneration. Tumorigenic capacity generally is limited to the nitrosamines. However, alkylamines may be converted to nitrosamines in the GI tract. General toxicological properties of alkylamines include the following:

- Irritancy increases up to six carbons and then decreases with further increasing molecular weight and loss of volatility.
- Unsaturated congeners are well absorbed following dermal exposure, while saturated congeners are not.
- Direct irritant potential is not affected by other functional groups, though sensitization potential may be altered.
- Aliphatic amine salts typically are weaker irritants than the parent molecule.

As noted, amines are strong irritants and represent a greater potential handling hazard than many other chemicals discussed in this chapter. Strong irritant properties stem from the fact that the amine portion of the molecule is a very corrosive functional group. Skin has some ability to resist changes in pH and can withstand chemical attack in the pH range of 2–10 for short periods without significant damage. However, if the acid has a pH lower than 2, or if base strength (pK) is much above 10, significant skin injury may occur quickly on dermal contact, as shown in Table 18.8.

Table 18.9 summarizes important dose–response relationships among the amines based upon molecular size, with

TABLE 18.8 Relative Base Strength of the Amines

Amine Compound	Base Strength (pK)
Methylamine	10.6
Dimethylamine	10.6
Trimethylamine	10.7
Ethylamine	10.8
Diethylamine	11.0
Triethylamine	10.7
Propylamine	10.6
Butylamine	10.6
Allylamine	9.5
Cyclohexylamine	10.5

TABLE 18.9 Relative Toxicity of the Amines

Amine	Oral LD_{50} (mg/kg)	Skin LD_{50} (mg/kg)	Dermal Effects
Methylamine	0.02	0.04	Necrosis
Ethylamine	0.4	0.4	Necrosis
Propylamine	0.4	0.4	Necrosis
Butylamine	0.5	0.5	Necrosis
Hexylamine	0.7	0.4	Slight necrosis

LD50 is the dose lethal to 50% of test animals; necrosis refers to cell/tissue death.

methylamine being dramatically more toxic by oral and dermal exposure than other simple amines.

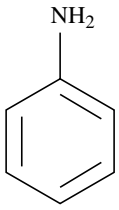
Because of penetrating and corrosive characteristics, amines are toxic to all tissues if absorbed in measurable amounts and can adversely affect a number of organs. Some systemic effects observed for lethal exposures are lung edema and hemorrhage, necrosis of the liver, necrosis/nephritis in the kidneys, and cardiac muscular degeneration. Two other common characteristics observed in amines are methemoglobin formation in the red blood cells (see Chapter 5) and sensitization to the chemical itself. Sensitization to amines is thought to occur because they bind to cellular proteins to form haptens, molecules recognized by the immune system. The body produces antibodies against haptens, and subsequent chemical exposure may cause allergic reactions. During this antibody–hapten response, the body releases histamine, which in turn induces arterial vasoconstriction, capillary dilation, fall in blood pressure, itching, and bronchoconstriction. These effects explain severe allergic responses: labored/difficult breathing, fainting, reddening and irritation where contact with the skin occurs, and possible anaphylactic shock. An uncommon feature of alkylamines involves the ability to simulate action of epinephrine (or adrenaline), an important neurotransmitter. Thus, it is easy to understand why some of the alkylamines may be able to mimic some epinephrine physiologic responses. Following are some of the general conclusions about alkylamine-induced sympathomimetic activity:

- Activity increases with the size of the alkyl chain, up to six carbons.
- For alkylamines with more than six carbons, heart rate decreases, and blood vessel dilation is more common.
- A branching of the alkyl carbon chain decreases activity of the chemical.
- Blood vessel constriction (pressor activity) varies with structure.
- Repeated exposure may cause cardiac depression and vasodilation.
- Convulsions may cause mortality at high acute exposures.

Human exposure to amine compounds should be avoided, since a number of these chemicals are regulated as carcinogens. In particular, this includes single- and multiple-ring compounds such as the dye components benzidine, *b*-naphthylamine, and 4-aminobiphenyl, all of which may induce bladder tumors. Other amine compounds may exhibit carcinogenic potential as well, including the nitrosamines. Nitrosamines, potential liver carcinogens, represent an interesting human risk issue, because alkylamines (including those generated during the digestion of food) theoretically may be converted to a nitrosamine by chemical conditions in the GI tract. To date, animal studies have failed to demonstrate this to be a significant problem in humans.

Aromatic Amines

The aromatic amines are ring hydrocarbons in which at least one hydrogen atom has been replaced by an amino group. They are synthesized by nitration of the aromatic hydrocarbon with subsequent reduction to the amine or by the reaction of ammonia with a chloro- or hydroxyhydrocarbon. Most aromatic nitro compounds are fat soluble and water insoluble such that they readily penetrate the skin and can quickly be absorbed via the lungs into the blood, immediately becoming systemic in their action. The most dominant toxic effects of the aromatic amine compounds are methemoglobin formation and cancer of the bladder and urinary tract. Sensitization is greater in aromatic amines compared to aliphatic amines. Other toxic effects include hematuria, cystitis, anemia, and skin sensitization. It is thought that the aromatic amines act indirectly on methemoglobin formation through their metabolites. Aromatic amino compounds are important as intermediates in the manufacture of dyes and pigments and are also used in the chemical, textile, rubber, and paper industries. Some aromatic nitro compounds, such as trinitrotoluene and pyridine, are used in manufacture of chemical explosives. Pyrrolidine, morpholine, aniline, pyridine, and similar compounds are used in crop protection as herbicides, fungicides, and insecticides.



Aniline

Aniline is a parent substance in the chemical industry for the synthesis of many compounds, including dyes, rubber accelerators, antioxidants, drugs, photographic chemicals, isocyanates, herbicides, and fungicides. Exposure is most likely to occur in occupational settings for those working with aniline. People living near contaminated waste sites may experience exposure through food and drinking water.

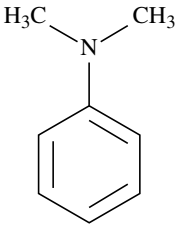
The most common symptoms of acute aniline intoxication are cyanosis, lacrimation, tremors, tachypnea, lethargy, methemoglobinemia, sulfhemoglobinemia, and Heinz body formation. Methemoglobinemia, which compromises the oxygen-carrying capacity of blood, is characteristic of aniline exposure. Because of its fat solubility, aniline can penetrate intact skin, and it is absorbed easily via inhalation. Moderate exposure to aniline by any route may cause cyanosis, but headache, weakness, irritability, drowsiness, dyspnea, and unconsciousness may occur following oxygen deficiency. Aniline ingestion can cause splenic enlargement, hemosiderosis, as well as marrow hyperplasia. Chronic

aniline poisoning at low levels is controversial; however, some CNS symptoms have been reported. Upper respiratory tract irritation and congestion have been observed following high levels of acute exposure. The TLV set by the ACGIH is 2 ppm, and the OSHA PEL is 5 ppm. Both of these precautionary values carry a "skin" notation, acknowledging significant dermal absorption potential.

CONCENTRATION PROFILE FOR ANILINE IN HUMANS

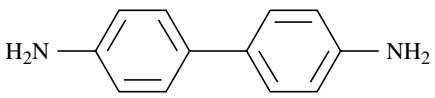
Concentration (ppm)	Endpoint
2	TLV
1.1	Odor threshold
7–53	Early symptoms of methemoglobinemia
100–160	Dyspnea, tachycardia, dizziness, headache

Carcinogenicity of aniline to humans is not clear, since it was mildly carcinogenic to rats, but not to mice. This may be due to interspecies metabolic differences. Industrially important exposures to aniline are evaluated by measurements of urinary aniline, aniline released from hemoglobin in the blood, or urinary *p*-aminophenol.



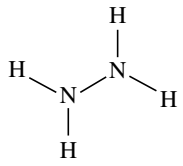
N, *N*-Dimethylaniline

N,*N*-Dimethylaniline (DMA) is used in the synthesis of dyestuffs, as a solvent, and as an analytical reagent. The compound can enter the body by inhalation, skin absorption, ingestion, and eye and skin contact. As with other amines, signs of intoxication in man include headache, cyanosis, dizziness, labored breathing, paralysis, and convulsions. Like aniline, DMA is readily absorbed through the skin, with the potential to produce methemoglobinemia. Few reports of industrial exposure are available with which to estimate the hazards of DMA. However, cases of severe exposure are accompanied by visual disturbances and severe abdominal pain. No evidence of reproductive or genetic toxicity is available for this compound. The TLV and PEL both have been set at 5 ppm by the ACGIH and OSHA, respectively.



Benzidine

Benzidine principally is used in the synthesis of dyes but also may be used as a hardener for rubber and as a laboratory reagent. A high incidence of bladder cancer has been reported among workers exposed on a chronic basis to benzidine. Inhalation and skin absorption are significant routes of benzidine exposure, which also may result in local contact dermatitis. Benzidine is classified as a known human carcinogen by a variety of agencies and exhibits mutagenic potential in numerous tests. No numerical exposure criteria have been set for benzidine because of the high occupational cancer potential and associated recommendation of no contact by any route of exposure. The skin notation has been given to benzidine to indicate a danger of cutaneous absorption.



Hydrazine

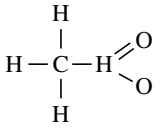
Due to its extreme reducing properties, hydrazine is used in many industrial chemical syntheses, photographic processes, and metallurgy, as well as a propellant. At sufficient concentration, hydrazine is a strong skin and mucous membrane irritant, convulsant, hepatotoxin, and moderate hemolytic agent. It is well absorbed through the lungs, GI tract, and intact skin, as shown by the fact that similar median lethal doses are observed in animal studies following oral, intravenous, and intraperitoneal administration. Effects noted after absorption by all routes include anorexia, weight loss, weakness, vomiting, excitement, and convulsions. Airborne exposure may produce eye and respiratory irritation, lung congestion, bronchitis, and pulmonary edema. Hydrazine is tumorigenic in mice, with lungs being the primary target organs. It is mutagenic in a variety of bacterial, plant, and mammalian test systems. The ACGIH TLV and OSHA PEL for hydrazine have been set at 0.01 and 1 ppm, respectively, based upon observations of lung, liver, and kidney damage at higher exposure levels. The skin notation has been given to hydrazine in recognition of cutaneous absorption.

Aliphatic and Alicyclic Nitro Compounds and Related Substances

Aliphatic and alicyclic nitro compounds are oily liquids exhibiting low solubility and low volatility. They are of limited industrial significance, but several members have important uses as specialized fuels and strong solvents. The nitroparaffins are acidic compounds and thus are rapidly neutralized with strong bases. They are easily chlorinated to form chloronitroparaffins (e.g., trichloronitromethane), which are moderate skin and mucous membrane irritants.

Saturated members of the class are not well absorbed, while unsaturated members exhibit significant dermal absorption.

Nitro derivatives of benzene and toluene also have prominent effects including sensitization, CNS depression, hepatotoxicity, and methemoglobinemia. Other reported problems include dermatitis, anemia, heart irregularities, and peripheral neuritis. Trinitrotoluene and dinitrobenzene compounds are well absorbed by all routes of exposure and can uncouple oxidative phosphorylation. Persons deficient in glucose-6-phosphate dehydrogenase are sensitive to hemolytic anemia. Some of these compounds are implicated in bladder tumors as well. Due to the diversity, specific toxicity of any member of this class should be reviewed individually prior to use.

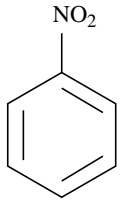


Nitromethane

Nitroparaffins, including nitromethane, are synthesized by vapor-phase nitration of propane. Nitromethane may be used as rocket propellant, a gasoline additive, solvent for cellulosic compounds, polymers and waxes, and in chemical synthesis. Nitromethane can produce narcosis, methemoglobinemia, mucous membrane irritation, CNS excitation, and liver damage. Based on periodic examinations of plant operators engaged in manufacture and handling of nitromethane for several years, chronic effects attributed to this compound are quite minor. Nitromethane does not exhibit significant mutagenic capacity in bacterial tests. Evidence of carcinogenicity has been demonstrated in several animal studies, resulting in classification as an A3 confirmed animal carcinogen with unknown relevance to humans by the ACGIH. Other agencies have classified nitromethane in the "possible" or "probable" categories of potential carcinogens. The TLV and PEL values that have been established by the ACGIH and OSHA for nitromethane are 20 and 100 ppm, respectively. A structural relative of nitromethane, 2-nitropropane, is used alone and in combination with other solvents. Its carcinogenic potential is classified similarly to nitromethane. The TLV and PEL values established by the ACGIH and OSHA for 2-nitropropane are 10 and 25 ppm, respectively.

Aromatic Nitro Compounds

The aromatic nitro compounds are formed by substituting a nitro group directly onto the benzene ring. Aromatic nitro compounds resemble other nitro compounds, producing dermal sensitization, CNS depression, and methemoglobinemia. These compounds have the capacity to uncouple oxidative phosphorylation and produce liver injury. Other reported adverse effects include dermatitis, anemia, cardiac irregularities, peripheral neuritis, and bladder tumors. Persons deficient in glucose-6-phosphate dehydrogenase are demonstrably sensitive to hemolytic anemia.

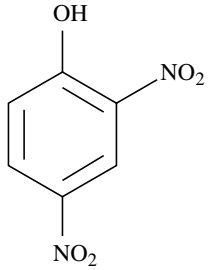


Nitrobenzene

Nitrobenzene is used as a solvent and intermediate in preparation of aniline, benzidine, and other chemicals and in shoe/metal polishes. The vapor is readily absorbed through the skin and lungs. Following absorption, nitrobenzene is metabolized to *p*-aminophenol, *p*-hydroxyacetanilide sulfate, *p*-nitrophenol sulfate, and *m*-nitrophenol sulfate. Ring hydroxylation and reduction are important steps in biotransformation of nitrobenzene, and metabolites are excreted mainly in the urine.

The primary nitrobenzene toxic effect involves methemoglobin formation, though the mechanism of action is not well understood. The primary thesis is that heme iron is oxidized by redox cycling of nitrobenzene metabolites. In subacute and chronic forms of poisoning, hemolysis causes anemia as a main symptom, but other symptoms such as headache, confusion, vertigo, nausea, spleen and liver damage, loss of cognition, hyperalgesia, paresthesia, and polyneuritis have been reported. The cyanogenic and anemiagenic potential of nitrobenzene are considerably greater than those of aniline, and potential for producing blood effects is second only to that of dinitrobenzene. There is some risk of reproductive toxicity, as judged by a fertility decrease following exposure to rats. Genotoxicity studies suggest that nitrobenzene is at worst weakly genotoxic.

Industrially important exposures to nitrobenzene may be evaluated by measurements of *p*-nitrophenol, expressed in conjunction with urinary creatinine. The TLV and PEL set by the ACGIH and OSHA are both at 1 ppm. The ACGIH has also given the skin notation to nitrobenzene. The U.S. EPA has concluded that nitrobenzene is likely to be a human carcinogen, while the ACGIH classifies the substance as an A3 confirmed animal carcinogen with unknown relevance to humans.



Dinitrophenol

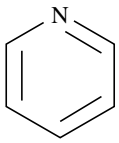
Among the isomers of dinitrophenol (DNP), the most common in industrial use is 2,4-DNP. The isomers often are used as a mixture and are involved in the synthesis of dyes, picric acid, picramic acid, herbicides, and the manufacture of photo developers.

Dermal application of DNP causes yellow staining of the skin and may cause dermatitis due to primary irritation or to allergic sensitivity. DNP disrupts oxidative phosphorylation, resulting in increased metabolism, oxygen consumption, and heat production. Acute poisoning is characterized by the onset of fatigue, sudden thirst, sweating, and chest pressure accompanied by rapid respiration, tachycardia, and a rise in body temperature. In less severe poisoning, symptoms may include nausea, vomiting, anorexia, weakness, dizziness, vertigo, headache, and sweating. Onset of effects is rapid, and death or recovery may occur within 1–2 days following massive exposure. Chronic exposure at lower levels may cause kidney or liver damage and cataract formation. No occupational standards have been set for DNP; however, an exposure limit of 0.2 mg/m^3 has been suggested based on dinitrocresol.

$$R-C\equiv N$$

Nitrile basic structure

Nitriles, including acrylonitrile and acetonitrile, are organic cyanides (i.e., C≡N compounds). They are nonpolar and readily absorbed by all routes. Some of these compounds dissociate to produce free cyanide, and adverse effects are consistent with cyanide poisoning. As noted previously for other organic solvents, unsaturated forms are more irritating than corresponding saturated analogs. The most common nitrile, acrylonitrile, is regulated as a suspected carcinogen by a number of occupational and environmental regulatory agencies, based primarily on data from animal studies. The TLV and PEL are both set at 2 ppm, with a skin notation.

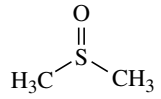


Pyridine basic structure

Pyridine is the parent compound for the pyridine series of substituted analogs. It is a flammable, unsaturated six-membered ring resembling benzene, but consisting of five carbons and one nitrogen, as opposed to six carbons. The compound exhibits a penetrating, nauseating odor that acts as a warning property. For most substituted benzene compounds, there is an analog in the pyridine series. Pyridine and its derivatives are used as solvents and raw materials in the manufacture of chemicals, explosives, paints, disinfectants, herbicides, insecticides, antihistamines, and vitamins. Pyridine exerts adverse effects on the CNS, GI tract, liver, and kidneys and has been used in epilepsy therapy. Skin irritation and photosensitization have been reported, and inhalation of vapor irritates membranes. Inhalation also causes anorexia, nausea/vomiting, gastric distress, headache, fatigue, faintness, and depression. Liver/kidney toxicity and death are reported at doses over 2 ml/day for 2 months. The TLV and PEL are set at 1 and 5 ppm, respectively. The ACGIH classifies pyridine as an A3 confirmed animal carcinogen with unknown relevance to humans.

The alkyl pyridines, and the parent molecules, are well absorbed from the GI tract, peritoneal cavity, lungs, and intact skin. The metabolic fate is not completely known, but hydroxylation, *N*-methylation, oxidation, and conjugation reactions occur. Elimination is rapid, limiting potential accumulation in tissues. Despite wide industrial application and limited medicinal use, reports of human poisoning are uncommon.

18.11 TOXIC PROPERTIES OF SULFUR-CONTAINING SOLVENTS



Dimethyl sulfoxide

Dimethyl sulfoxide (DMSO) is an industrial solvent that also has a wide applicability in the pharmaceutical area to solubilize water-insoluble medication. It has the ability to carry solutes into the skin from which they are slowly released into the blood and lymph. It is remarkably nontoxic, though some evidence suggests that DMSO has some hepatic and renal effects. DMSO has been reported in low concentrations in some food products, as well as in the atmosphere. It can inhibit enzyme reaction, but does not affect thyroid function. DMSO reportedly is metabolized to dimethyl sulfide and subsequently respired. DMSO lowers body temperature and influences toxic effects of aromatic hydrocarbons. It may influence preservation of leukemia cells and proved to be a potent inducer of erythroid differentiation in cultured erythroleukemic cells. DMSO is not mutagenic and is actually used as a solvent in mutagenicity tests. Teratogenic potential has been reported, related to route of administration, dose, and gestational period. It also has the ability to influence activity of materials with carcinogenic potential. No regulatory values are available for DMSO.

$$S=C=S$$

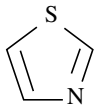
Carbon disulfide

Carbon disulfide is a flammable, toxic solvent that historically was used in manufacture of rayon, cellulose fibers, rubber vulcanizers, and pesticides, as well as carbon tetrachloride. Its previous use as a grain fumigant has been discontinued. While carbon disulfide is odorless, impurities may impart objectionable sulfur odors.

The principal toxicological effect at high air concentrations is narcosis, accompanied by headache, visual disturbance, respiratory disturbance, and GI effects. Ingestion of as little as 15 ml may be fatal. Aside from the neurological changes associated with acute exposure, the greatest toxicological concern for carbon disulfide relates to demonstrated ability to induce peripheral polyneuropathy and psychoses in some chronically exposed individuals. The latter effects reportedly resolve following exposure cessation, but the characteristic neuropathies may persist. These include reflex decrements in extremities, glove/stocking sensory loss, and decreased nerve conduction velocity. Carbon disulfide is not regulated as a carcinogen in an industrial or environmental context. The TLV and PEL set by the ACGIH and OSHA are 1 and 20 ppm, respectively. A skin notation is given to carbon disulfide by the ACGIH. Industrially important exposures to carbon disulfide may be evaluated by measurement of 2-thiothiazolidine-4-carboxylic acid (TTCA), expressed in terms of urinary creatinine.

CONCENTRATION PROFILE FOR CARBON DISULFIDE IN HUMANS

Concentration (ppm)	Endpoint
1	TLV
0.1	Odor threshold
150–300	Polyneuropathy
5000	Acute lethality



Thiazole basic structure

Thiazoles, in particular benzothiazole and mercaptobenzothiazole, are used in the rubber vulcanizing process and as fungicides. Benzothiazole occurs in such a small quantity that it is not considered to be a major health threat, though it is moderately toxic. Mercaptobenzothiazole, when heated, may react with oxidizing material and emit toxic decomposition products. The main effect of exposure is allergic contact dermatitis, and the compound is considered a potent allergen. Subcutaneous tests on mice showed possible carcinogenic potential, though it is not regulated as such. No ACGIH or OSHA exposure values are available for mercaptobenzothiazole.

18.12 SUMMARY

As discussed in this chapter, the common toxicological effects attributed to individual solvents and related materials include:

- CNS depression and other neurotoxic effects
- Respiratory irritation
- Dermal effects, including irritation
- Nephrotoxicity
- Hepatotoxicity
- Carcinogenicity

The emphasis of the chapter reflects important chemical properties, behavior, and effects, citing appropriate examples. The chapter reviews ranges of toxic effects that may be expected from selected chemical classes, which should serve as a good introduction for other sources of detailed information. The following section provides valuable supplemental reference sources related to solvent chemistry and toxicology.

SUGGESTED READING

Agency for Toxic Substances and Disease Registry (ATSDR). *Toxicological Profiles*. Atlanta: ATSDR; 1988–2010.

American Conference of Governmental Industrial Hygienists (ACGIH). *Documentation of Threshold Limit Values and Biological Exposure Indices*. 7th ed. ACGIH; 2001–2014.

California Environmental Protection Agency Office of Environmental Health Hazard Assessment (CalEPA OEHHA). (1999). Part 1: Determination of Acute Reference Exposure Levels for Airborne Toxicants. In Air Toxicology and Epidemiology Section.

International Agency for Research on Cancer (IARC). *Monographs*. Lyon: World Health Organization; 1972-present.

National Institute of Occupational Safety and Health (NIOSH). *Pocket Guide to Chemical Hazards*. Cincinnati (OH): US Department of Health and Human Services, Centers for Disease Control and Prevention; 2006.

National Toxicology Program (NTP). *NTP-CERHR Monograph on the Potential Human Reproductive and Developmental Effects of Bisphenol A*. Research Triangle Park (NC): US Department of Health and Human Services; 2008.

National Toxicology Program (NTP). *Report on Carcinogens*. 12th ed. Research Triangle Park: U.S. Department of Health and Human Services; 2011.

United States Environmental Protection Agency (USEPA). Chemical Specific Toxicological Reviews. In support of summary information on the Integrated Risk Information System; 1998–2014.

United States Environmental Protection Agency (USEPA). Guidelines for Carcinogen Risk Assessment. Risk Assessment Forum. Washington (DC): U.S. Environmental Protection Agency; 2005.

United States Environmental Protection Agency (USEPA). 2007. Technology Transfer Network Air Toxics Web Site. Available at http://www.epa.gov/ttnatw01/hlthef/formalde.html. Accessed August 05, 2014.

United States Environmental Protection Agency (USEPA). *Toxicological Review of Trichloroethylene*. Washington (DC): U.S. Environmental Protection Agency; 2011a.

United States Environmental Protection Agency (USEPA). 2011b. National Priorities List. Web Site. Available at http://www.epa.gov/superfund/sites/npl/. Accessed May 2011.

United States Environmental Protection Agency (USEPA). *Toxicological Review of Tetrachloroethylene*. Washington (DC): U.S. Environmental Protection Agency; 2012.

19

NANOTOXICOLOGY

HONGBO MA, STEVE DIAMOND, GEORGIA HINKLEY, AND STEPHEN M. ROBERTS

Nanotechnology represents one of the most promising technologies of the twenty-first century and has been considered to be a new industrial revolution. It is defined as the development and application of materials and structures with nanoscale dimensions, usually in the range 1–100 nm. The prefix "nano" is derived from the Greek word "nanos" meaning "dwarf." Nanomaterials refer to any materials with at least one dimension in the range of 1–100 nm ($1\,nm = 10^{-9}\,nm$). Particles in the nanosized range have been present on earth for millions of years. Soot, for instance, a product of incomplete combustion of fossil fuels and vegetation, has a particle size in the nano- to micrometer range and therefore falls partially within the "nanoparticle" domain. Only recently nanoparticles have attracted great attention because of the increasing manufacturing and manipulation of such materials by humankind. Therefore, the term "nanoparticles" usually applies only to engineered particles but not particles under 100 nm that occur naturally or are by-products of other processes such as welding fumes, fire smoke, or carbon black.

Nanomaterials often display novel properties as compared to their bulk counterparts, such as enhanced magnetic, catalytic, optical, electrical, and mechanical properties. For example, inert elements like platinum or gold become excellent chemical catalysts when in nanoscale. Normally stable aluminum becomes combustible and copper becomes transparent. Such novel properties of nanomaterials have rendered their increasing application in a broad range of areas, from industry products such as tires, catalysts, electronic component, window sprays, paints, and coatings to commercial products such as sporting goods and personal care products and medical applications such as antimicrobial agents, clinical diagnosis, imaging, and drug delivery. A report by the U.S. National Science Foundation in 2006 estimated that the global market for nanotechnology may reach $1 trillion or more within 20 years. The novel properties of nanomaterials, and the challenges in evaluating their toxicity, have led to the development of nanotoxicology as a new field.

This chapter will discuss:

- Classifications and uses of nanomaterials
- Special considerations and challenges in assessing the toxicity of nanomaterials
- The toxicity of selected nanomaterials
- Environmental fate and effects of nanomaterials
- Mechanisms of toxicity

19.1 CLASSIFICATION OF NANOMATERIALS AND THEIR POTENTIAL APPLICATIONS

Nanomaterials are broadly defined by sizes. In theory, all conventional materials like metals, semiconductors, glass, ceramic, or polymers can be engineered with a nanoscale dimension. Therefore, there are different approaches for classification of nanomaterials. The most widely accepted classification is based on chemical composition, from which nanomaterials can be classified into four main categories: carbon-based nanomaterials, metal/metal oxide nanomaterials, quantum dots (QDs), and dendrimers.

Carbon-based nanomaterials include fullerene, carbon nanotubes (CNTs), and related products. Manufacturing of all these nanomaterials originates from the discovery in 1985 of fullerenes (also known as buckyballs), a 60-carbon atom

Principles of Toxicology: Environmental and Industrial Applications, Third Edition. Edited by Stephen M. Roberts, Robert C. James and Phillip L. Williams.

hollow sphere. Fullerenes and CNTs are produced in quantities as high as 1500 t/year by companies such as Frontier Carbon Corporation in Tokyo, Japan, and Fullerene International Corporation in Tucson, AZ, United States. Due to their excellent thermal and electrical conductivity, these materials are widely used in plastics, catalysts, battery and fuel cell electrodes, sensors, water purification systems, and automotive industries.

Several varieties of organic nanoparticles have been engineered by the pharmaceutical industry for drug delivery, providing a variety of advantages compared to conventional systems. Nanoparticle drug delivery is particularly advantageous for acid-sensitive oral medications such as protein or RNA-based therapies. These systems provide safe passage through the harsh gastric environment and deliver therapeutics to absorptive surfaces in the small intestine, often designed to degrade in a neutral pH solution. Other routes of administration with certain applications can also greatly benefit from using nanoformulated delivery systems. For example, the harsh side effects of chemotherapy drugs can be lessened by nanodelivery systems that are coupled to a targeting agent designed to release medications at the site of a tumor. Liposomes, solid lipid nanoparticles, polymeric micelles, mesoporous silica particles, and nanocrystals are just a few examples of engineered nanoparticles used in the pharmaceutical industry.

Metal-containing nanomaterials have received considerable attention and have a broad range of applications. Titanium dioxide (TiO_2) and zinc oxide (ZnO) are widely exploited for their photocatalytic properties in industrial applications such as solar cells, paints, and coatings. They are also finding extensive applications in sunscreens and cosmetics due to their excellent UV-blocking capability and visible transparency at nanoscale. Production of these metal oxides for use in skincare products is estimated to be 1000 t/year in 2005–2010. Cerium dioxide (CeO_2) is mainly used as a combustion catalyst in diesel fuels to improve combustion efficiency and control emissions. The most widely used nanoparticulate zerovalent metals include silver (Ag), iron (Fe), and gold (Au). Nanosilver has been incorporated into the greatest number of consumer product applications due to its excellent antimicrobial and antiviral activities. Nanoscale iron is most widely used in environment remediation of waters, sediments, and soils. Nanoscale gold has been used in medical applications as vectors in tumor therapy.

QDs are semiconductor nanocrystals that have extraordinary optical and electrical properties due to the quantum-confined nature of their energy levels. They are usually composed of a reactive core (such as cadmium selenide [CdSe], cadmium telluride [CdTe], or zinc selenide [ZnSe]) that controls its optical properties and a protective shell (such as silica or ZnS) that protects the core from oxidation and enhances the photoluminescence yield. QDs are used largely in medical applications such as diagnostic imaging or targeted therapeutics. Extended applications of these QDs include solar cells and photovoltaics, security inks, photonics, and telecommunications. Dendrimers are multifunctional polymers whose size, topology, flexibility, and molecular weight can be controlled. Their applications typically involve conjugating other chemical species to their surface, which can function as detecting agents, affinity ligands, targeting components, imaging agents, or pharmaceutically active compounds.

Nanomaterials are also being used as direct food additives and in food packaging materials as antibacterial agents. Nanosized silica and titanium dioxide are both used in a number of food products—silica as an anticaking and thickening agent for dry food products and titanium dioxide as a whitening agent in many candies and gums. Sources estimate that the current human exposure to nanosized silica and titanium dioxide from food exceeds 40 mg per person per day. In addition, silver and zinc nanoparticles are being used in food packaging, for example, absorbent meat pads, as antibacterial agents to extend the shelf life of foods. Release of embedded particles from packaging materials may lead to further human exposure to nanomaterials from foods.

With increasing use in commercial and consumer products, both intentional and unintentional exposures to nanomaterials and their release into the environment are inevitable. Unintentional release includes atmospheric emissions and solid or liquid waste discharge during manufacturing processes and intentional release includes their uses in remediation of environmental contamination, as well as in consumer products such as paints, fabrics, and personal care products. In some cases, disposed nanoparticles will ultimately be deposited on land and surface water systems and have the potential to interact with biota. Atmospheric emissions (e.g., diesel fumes) may lead to inhalation exposure in humans and animals, and any waterborne contamination is also of concern for dermal and ingestion exposure. Certain synthetic nanoparticles have been detected in the environment, and modeled concentration of engineered nanomaterials can reach μg/l levels (Table 19.1). The very same properties that lead to the technical advancement in nanotechnology may also lead to unexpected and unanticipated consequences upon interaction with biological systems. For example, the small sizes of nanomaterials impart a different biokinetic behavior of the materials and enable them to reach more distal regions of the body compared to their bulk counterparts. To fully recognize the benefits of nanotechnology and to ensure a sustainable development and application of it, it is essential to understand its implications on human and environmental health at an early stage. Concerns have been raised from both scientific community and general public regarding the environmental health and safety of nanotechnology.

TABLE 19.1 Predicted Environmental Concentrations of Engineered Nanomaterials (ENMs) in the Environment in Europe and the United States

	Surface Water (μg/l)	STP Effluent (μg/l)	Sediment (Δμg/kg/year)	Soil (Δμg/kg/year)
	TiO_2			
Europe	0.015 (0.012, 0.057)	3.47 (2.50, 10.80)	358 (273, 1409)	1.28 (1.01, 4.45)
United States	0.002 (0.002, 0.010)	1.75 (1.37, 6.70)	53 (44, 251)	0.53 (0.43, 2.13)
	ZnO			
Europe	0.010 (0.008, 0.055)	0.432 (0.340, 1.42)	2.90 (2.65, 51.7)	0.093 (0.085, 0.661)
United States	0.001 (0.001, 0.003)	0.3 (0.22, 0.74)	0.51 (0.49, 8.36)	0.050 (0.041, 0.274)
	CNT			
Europe	0.004 (0.003, 0.021)	14.8 (11.4, 31.5)	241 (215, 1321)	1.51 (1.07, 3.22)
United States	0.001 (0.001, 0.004)	8.6 (6.6, 18.4)	46 (40, 229)	0.56 (0.43, 1.34)
	Ag			
Europe	0.764 (0.588, 2.16)	42.5 (32.9, 111)	952 (978, 8593)	22.7 (17.4, 58.7)
United States	0.116 (0.088, 0.428)	21.0 (16.4, 74.7)	195 (153, 1638)	8.3 (6.6, 29.8)
	Fullerenes			
Europe	0.017 (0.015, 0.12)	5.2 (4.23, 26.4)	17.1 (6.22, 530)	0.058 (0.057, 0.605)
United States	0.003 (0.002, 0.021)	4.6 (4.49, 32.66)	2.5 (1.05, 91.3)	0.024 (0.024, 0.292)

Source: Adapted with permission from Gottschalk et al. (2009). © 2009 American Chemical Society.
Numbers in parentheses are the lower and upper quantiles ($Q_{0.15}$ and $Q_{0.85}$). For surface and sewage treatment plant (STP) effluents, the numbers represent 2008 ENM concentrations; for sediment and soil, the numbers illustrate the annual increase of ENM concentrations.

19.2 HOW IS NANOTOXICOLOGY DIFFERENT?

Nanotoxicology is the study of toxic effects of nanomaterials. It is proposed as a new branch of toxicology to address the gaps in knowledge concerning the toxicological effects caused by nanomaterials. It encompasses the physicochemical properties of nanomaterials, routes of exposure, toxicokinetics and toxicodynamics, and regulatory aspects. It also involves proposing and developing reliable and robust test protocols for nanomaterials in human and environmental risk assessment. It should be pointed out that naturally occurring nanoparticles, such as nanosized metal oxides, exist in all ecosystems and the living organisms have adapted to them during evolution. However, the potential harmful properties of synthetic nanomaterials have to be evaluated as toxicological data on these materials are just emerging.

Nanoscale materials behave differently than their macroscale counterparts because of their smaller size. Below about 100 nm, the behavior of elements follows the rules of quantum physics. Gravity is not important and van der Waals forces govern the attraction/repulsion behavior between particles. Surface charge becomes very important and agglomeration of particles occurs readily, depending on environmental conditions. Some nanoparticles can easily penetrate cell walls and membranes without the requirement of phagocytosis for entering cells. Toxicity or behaviors that are surface related are exacerbated by the increasing surface-to-volume ratios at the nanoscale. These novel and unique properties of engineered nanomaterials bring special challenges to nanotoxicology as compared to conventional toxicology studies.

Importance of Characterization

Characterization of the physicochemical properties of nanomaterials is critical because these are usually the determinants of biological activity and effects of the nanomaterials. Characterization of nanomaterials is much more complex and indispensable as compared to their larger counterparts (which are usually confined to chemical composition and purity determination), because nanomaterials can differ widely in terms of particle size, shape, surface area, surface chemistry, crystallinity, solubility, agglomeration state, and chemical composition.

There has been a consensus within the scientific community that thorough characterization of nanomaterials is essential during toxicological studies, although disagreement exists regarding the key "essential" characterization information to be obtained. These properties may include, but are not limited to, aggregation/agglomeration, size, dissolution (solubility), surface area, surface charge, and chemical composition. In addition, some key abiotic factors in the aqueous media, such as pH, presence of natural organic matter (NOM), and ionic strength, should also be measured. In practice, complete characterization is not always possible due to extensive instrumentation and specialized technical expertise required; therefore, the extent of characterization usually depends on objectives of a specific study and prioritizing characterization measurements on the basis of their known or suspected influence on toxicity.

A variety of methods are available for characterizing nanoparticles. These include high-resolution imaging techniques such as scanning and transmission electron microscopy (SEM/TEM) that can be used to measure the shape,

size, and crystalline nature of the particles. Size is usually the most important criterion that differentiates a nanoparticle for functionality among all attributes of the nanomaterial. Toxicity data is meaningless without proper determination of size and/or size distribution of the test material. Nanoparticles tend to agglomerate in environmental or biological fluids. Agglomeration involves adhesion of particles to each other mainly due to van der Waal's forces that become quite prominent at nanoscale due to increased surface area-to-volume ratios. The rate and extent of agglomeration may vary based on interaction with macromolecules and biological fluids once the nanoparticles are introduced inside living systems. A nanoparticle suspension is a dynamic system as agglomeration or dissolution is ongoing and agglomeration is highly dependent on physiochemistry of the solution. Agglomeration behavior is expected to affect nanotoxicity as it changes the size, surface area, and sedimentation properties of the nanoparticles and should be considered for toxicity studies.

The most commonly employed techniques for particle size determination are Brunauer–Emmett–Teller (BET), dynamic light scattering (DLS), and TEM, although a variety of methods are available. The BET method is typically used for the calculation of surface area of solids by physical adsorption of gas molecules on the solid surface. TEM gives direct measurement of particle size, size distribution, and morphology by image analysis. DLS has the ability to provide the size and size distribution information in relevant liquid media used for exposure. Each technique has its own advantages and disadvantages and no single technique is appropriate for all materials. Techniques for nanoparticle surface analysis include electron spectroscopies, ion-based methods and low-energy ion scattering, and scanning probe microscopy (including atomic force microscopy and scanning probe microscopy). Solubility of nanoparticles can be measured by ultrafiltration, dialysis, and flow field fractionation coupled to inductively coupled plasma mass spectrometry (ICP-MS). Adequate characterization of nanoparticles often involves a variety of complementary analysis methods. However, none of these methods are well suited to quantifying the amount of nanoparticles in biological (tissues and fluids such as blood) or environmental media (soil, sediment, or water), which remains an area of high research need.

Issues with Test Materials

Another unique challenge facing nanotoxicology is that there are large numbers of materials to be tested and the numbers keep increasing as new materials are developed and produced. As it is not feasible to conduct toxicity test for every nanomaterial, it is highly desirable for the scientific community to have access to well-characterized nanomaterials with known toxicity and preferably a well-documented toxicity mechanism that may be used as reference materials for hazard assessment. Reference materials play an important role in particle toxicology as particles may have variable toxicity depending on type, origin, chemistry, size distribution, etc. Reference materials have been used historically in particle toxicity studies. For example, asbestos fibers from five different locations in the world and with specific physicochemical properties are prepared by the Union of International Cancer Control as reference materials for toxicity study of asbestos.

Development of reference materials for nanomaterial toxicity testing is just beginning. A set of nanoparticle reference standards consisting of colloidal gold with nominal diameters of 10, 30, and 60 nm was released by National Cancer Institute in 2007. A number of international organizations and initiatives are underway to develop reference materials for nanotoxicology studies. For instance, the Organisation for Economic Co-operation and Development (OECD) has produced a list that comprises fullerenes, single- and multiwalled CNTs, carbon black, polystyrene, dendrimers, nanoclays, and nanoparticles of Ag, Fe, TiO_2, Al_2O_3, CeO_2, ZnO, and SiO_2. All these materials are already in commercial production and applications.

Other common difficulties with test materials include consistency of the materials and quantities available for study. As manufacturing processes for new nanomaterials are being worked out, the nanomaterials can vary substantially from lot to lot with respect to key characteristics. This reinforces the need to characterize each batch or lot of test material received to insure that it meets the intended specifications. Also, new nanomaterials may be available only in very small quantities (e.g., milligram amounts). This can make obtaining quantities necessary to perform toxicity studies, especially *in vivo* studies, difficult.

Issues with Dosimetry

Accurate assessment of toxicity relies on correct assessment of dose–response relationships. The dose metric is one of the critical factors that determine the shape of dose–response curves. In general, toxicologists express doses by mass. However, this may not be the best approach for nanomaterials given their extreme low mass and high surface area-to-volume ratio. Size and surface area are two important properties from a toxicological perspective. As size decreases, the proportion of surface atoms increases making the particle surface more reactive. Such enhancement of surface reactivity at nanoscale is usually responsible for the nanoparticle-elicited biological response. Therefore, it is generally believed that particle surface area is a more appropriate dose metric than the traditional mass-based measure of dose. For example, particle BET surface area has been suggested as the most appropriate parameter to evaluate the inflammatory potential of ultrafine carbon particles in mice

bronchoalveolar lavage. Similar conclusions have been reached for titanium dioxide nanoparticles in terms of their capacity for reactive oxygen species (ROS) generation. This is consistent with the mechanistic consideration that cellular response may be related to surface area of a nanoparticle due to effects arising from particle surface chemistry/reactivity.

Particle number represents another candidate dose metric. This is especially true when the toxic responses of nanoparticles are a result of obstruction of cellular processes by the physical presence of poorly soluble particles. The number of particles in the exposure medium has been considered to be the best parameter to describe the toxicological effects of different types of nanoparticles on acute lung inflammation in rats and mice after a comparison of four types of dose metrics: particle number, joint length (product of particle number and mean size), BET surface area, and calculated particle size based on surface area.

The complexity of particle properties and mechanisms of action preclude generalization of the appropriate dosimetry for all nanoparticles. Surface area, particle number, or even mass per volume could be the relevant dose metric in risk assessment of a given material, depending on the specific test conditions and the biological effects of interest. As the variability in results of similar nanotoxicity studies will continue to exist, it is necessary to analyze the results obtained with respect to different types of dose metrics so that a convincing conclusion regarding the comparative toxicity of a nanomaterial can be reached. This in turn will help in deciding the suitability of different dose metrics for nanotoxicological studies.

It should be noted that even selection of a suitable dose metric for *in vitro* nanotoxicity test is not enough for deciding cellular dose without considering other factors associated with nanoparticles. Unlike soluble chemicals, particles can diffuse, agglomerate, and settle depending on their physicochemical properties and media characteristics. Therefore, the dose introduced to the exposure system does not necessarily reflect the dose reaching the biological target. Yet in most *in vitro* nanotoxicity studies, it is the administered dose or media concentration that is measured as done in classical toxicology.

Suitability of Standard Toxicity Tests

In addition to considering dosimetry issues when designing a nanotoxicology study, it is important to evaluate the appropriateness of toxicity assays for the materials being tested. Standard toxicity assays were designed and validated for soluble chemical materials. Particulate matter presents a unique situation because these materials may possess physiochemical properties that are able to interfere with standard assays. Due to their high surface area-to-volume ratio, particles are capable of adsorbing large quantities of proteins and many particle types also have surface catalytic properties. Both of these attributes may cause assay interference. In addition to direct interaction with assay components, particles have the potential to interfere with endpoint measurement through spectral properties. Metallic particles have significant potential for interference via light absorption and fluorescent particles via light emission and/or absorption. It is important to consider and control for the potential of assay interference when performing standard toxicity assays with nanomaterials.

19.3 HUMAN NANOTOXICOLOGY

Bioavailability of Nanomaterials

Human exposure to nanomaterials is likely to occur incidentally as well as intentionally via several routes including inhalation, ingestion, and dermal exposure. Inhalation exposure to particulate matter has been studied for quite some time and is most well understood. The most important aspect for inhalation bioavailability is particle size—particles larger than 10 μm are trapped in the nose and upper airways, while particles less than 10 μm are able to deposit much lower in the respiratory tract in a size-dependent manner. Particles with a diameter less than 250 nm are able to translocate into the bloodstream and achieve systemic circulation. Similarly to inhalation exposure, ingestion of nanomaterials leading to systemic exposure is also largely size dependent. Particles greater than 3 μm are generally not well absorbed in the gastrointestinal tract (GIT), while particles near 1 μm in diameter are available for immune-mediated uptake via specialized intestinal immune centers called Peyer's patches. Smaller particles can achieve uptake via enterocytes, either passively through cell turnover or actively through receptor-mediated uptake. However, it should be noted that overall oral bioavailability of most particle types is quite low. Dermal absorption of intact particles through noncompromised skin is almost nonexistent.

For all routes of exposure, there are two very important caveats to consider that can greatly alter bioavailability: dissolution and agglomeration. As size is an important aspect for systemic uptake following all three routes of exposure, processes that affect particle size are always important to evaluate. Dissolution in the pulmonary and gastric environments can greatly increase overall bioavailability, by both decreasing average particle size and through the uptake of ions. Dissolution can also greatly increase dermal bioavailability; however, dermal exposure presents a special situation due to the potential for UV interaction with deposited particles. Breakdown of particles by UV exposure can even further increase dermal bioavailability. In contrast to dissolution, agglomeration in high-salt biological environments will have the opposite effect for all routes of exposure, resulting in decreased bioavailability. In addition to particle

properties that are likely to influence bioavailability, any compromise to the barrier integrity of the GIT or skin is likely to increase bioavailability and should be considered.

Toxicity of Nanomaterials

As mentioned earlier, the toxicology of inhaled particulate matter has been studied for centuries but has more recently been expanded to include a category for nanosized or "ultrafine" particles. Inhaled particles that achieve alveolar deposition have the potential to cause inflammation, leading to several diseases including asthma, fibrosis, chronic obstructive pulmonary disease, and cancer. Epidemiological analysis of occupational exposure to various metal and organic dusts has been shown to have an association with the development of fibrosis and cancer; however, these associations are confounded by exposure to a mixture of particles from a wide range of sizes. Rodent studies have shown that subchronic and chronic exposure to silver, titanium dioxide, zinc oxide, and copper oxide particles causes an inflammatory response in the lung. Inhalation exposure to nanosized fibers like CNTs has been shown to cause fibrotic-like responses, similar to those observed in asbestos-exposed humans.

In general, much less toxicity is observed for ingested nanomaterials. The GIT tissue is protected from direct toxicity by a thick mucin layer, and systemic toxicity is limited due to low overall bioavailability. Subacute exposure to zinc oxide in rats has been shown to cause changes in clotting times and dose-dependent increases in aspartate aminotransferase (AST) and alanine aminotransferase (ALT) levels, and subchronic exposure to silver in rats causes dose-dependent increases in alkaline phosphatase (ALP). In addition to direct tissue and systemic toxicity, disruption of GIT microbiota is a potential toxic response to ingestion of nanomaterials due to the antibacterial properties of many metal particles.

Other than the case of beryllium, systemic toxicity due to dermal nanoparticle exposure is quite rare. Due to the known reactions of the bulk materials, metal particles such as nickel and cobalt have the potential to cause allergic-type reactions, such as contact dermatitis. Beryllium presents a special case in which a lung disorder known as chronic beryllium disease can be initiated by dermal exposure to beryllium dust. This disorder requires both a sensitization phase, achieved by inhalation or dermal exposure to beryllium, followed by a progression phase of beryllium inhalation. Like most occupational exposures, these associations are confounded by exposure to a large range of particle sizes, but it can be expected that a portion of the dust generated during beryllium processing is in the nanorange.

When considering the observed toxicity of nanomaterials, it is important to consider the role of dissolution, particularly for metal particle exposures. Many metals in the ionic form are known to be toxic (e.g., copper) and may account for a portion or all of observed toxicity. Especially following ingestion, when particles can be exposed to the harsh pH conditions of the stomach for more than 2 h, dissolution should be carefully considered and measured when possible.

19.4 ENVIRONMENTAL NANOTOXICOLOGY

Environmental Fate and Transport of Nanomaterials

Nanomaterials may be released to the air, soil, or water both unintentionally and intentionally. Nanoparticles can be highly mobile in the environment due to their small size and light mass. Once entering into the environment, nanomaterials undergo abiotic interactions with a variety of environmental factors, leading to physical and chemical alternations of the materials. These alternations usually determine the transport and fate of nanomaterials and consequently their bioavailability and toxicity to environmental receptors. A logical chain of events leading to nanomaterial toxicity in the ecosystem is shown in Figure 19.1.

Physicochemical properties of nanoparticles have significant impact on their fate and transport in the environment. These parameters include chemical composition, surface area, size distribution, surface charge, surface coating, and solubility. Particle aggregation/agglomeration increases deposition rates of nanomaterials and thus reduces their mobility in the environment. Although there are good models to predict agglomeration and deposition of spherical particulates, such models are not yet available for behavior of nanoparticles that have nonspherical shape and complex surfaces. This makes it difficult to predict nanoparticle transport and fate in environmental media. Furthermore, nanoparticle behavior is highly dependent on water chemistry parameters such as temperature, pH, hardness, ionic strength, as well as presence or absence of NOM. For example, pH and temperature have great impact on solubilization of metal oxide nanoparticles such as ZnO. Ionic strength can affect particle aggregation/agglomeration dramatically. Salinity can also significantly increase agglomeration of nanoparticles. Most publications on environmental fate and toxicity of nanomaterials provide characteristics on pristine materials and are usually limited to particle size and chemical composition. Environmental matrices such as surface water and soil differ widely in pH and ionic composition; thus, agglomeration/aggregation and adsorption, and therefore particle behavior and mobility in the environment, are expected to be far more complex than observed in the laboratory.

Nanoparticles inevitably interact with other environmental contaminants and such interaction can either amplify or alleviate the toxicity of those compounds. Nanoparticles may act as transfer vectors of other contaminants in aquatic environments. In natural aquatic systems, soluble contaminants are predominantly bound to particle surfaces or

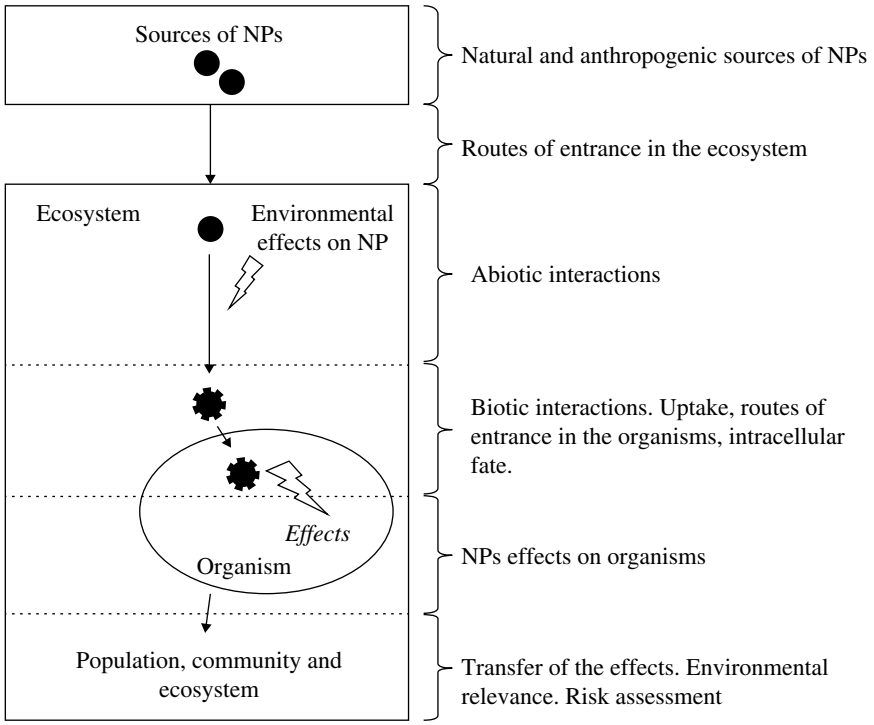


FIGURE 19.1 The logical chain of events accounting for the bioavailability and toxicity of nanomaterials starts with the nanomaterials' entrance into the ecosystem, proceeds by abiotic interactions between the nanomaterials and environmental factors, leads to bioavailability and toxicity to ecological receptors, and finally results in effects on population, community, and ecosystem levels. *Source:* Reprinted with permission from Navarro et al. (2008). © 2008 Springer.

complexed by humic substances. If the nanoparticles with the adsorbed pollutants are taken up by living organisms, a toxic effect could be a consequence, either induced by the nanoparticles, the pollutant, or both together. For example, the presence of C60 aggregates can increase the toxicity of phenanthrene for *D. magna* by 60% as the absorbed phenanthrene becomes bioavailable. It could also be that no effect is observed if the bound pollutant is not bioavailable and the nanoparticle itself is not toxic. Several studies have reported decreased bioavailability of organic pollutants in sediment and benthic invertebrates in the presence of black carbon, which is attributed to the strong sorption of the organic compounds to black carbon. On the other hand, such interaction between nanoparticles and other environmental pollutants can have advantageous applications. The unique structure and electronic properties of some nanoparticles can make them especially powerful adsorbents and can potentially be used for environmental remediation. A good example is zerovalent iron nanoparticles, which are widely used in groundwater remediation due to their excellent adsorption capacity and reducing power for both organic and inorganic pollutants.

Sediment is likely to be the final sink for nanoparticles in aquatic systems. The fate of nanoparticles in soil may be more complex and is largely unknown. The nanoparticles may bind chemically to soil particles or remain separate by residing either on the particle surface or in the pore space. They may adsorb organic matter that can stabilize particle surface charge and cause particle agglomeration. Studies on nanoparticle mobility in porous media, such as groundwater aquifers or sand filters, have shown that mobility is a function of surface chemistry and particle size. Nanoparticles intentionally released into the environment to remediate contaminated soils or sediments, such as zerovalent iron, have polymer coats to reduce agglomeration and increase water solubility.

Bioavailability and Toxicity of Nanomaterials in the Environment

Bioavailability is a dynamic concept that considers physical, chemical, and biological determinants of contaminant exposure and dose by incorporating concepts of environmental chemistry and toxicology and integrating contaminant concentration, fate and transport, and an organism's behavior in a given environment. Bioavailability of nanoparticles depends on the (i) physicochemical properties of the particles (aggregation, solubility), (ii) nanoparticle–organism interaction in the environment, and (iii) the target organism. Organisms living in environments containing nanoparticles will incorporate them into their bodies, through direct ingestion or entry across epithelial boundaries such as gills or

body wall. A consistent body of evidence shows that nanoparticles are taken up by a wide variety of environmental species and cause toxicity.

Aquatic Species Understanding toxic effects on the aquatic environment from nanomaterial exposure is important as it ultimately receives runoff and wastewater from domestic and industrial sources. Most of the currently available nanoecotoxicological data concerns crustaceans, followed by bacteria, algae, and fish. Crustaceans are the most numerous and ecologically important group of invertebrates in both freshwater and marine ecosystems. They also play an important role in regulatory toxicity testing as a part of the base set of organisms required for environmental risk assessment.

Most abundant data are available for those synthetic nanomaterials that are of current or expected applications in commercial products. These include TiO_2, C60, ZnO, Ag, CNTs, and CuO. Recent studies of TiO_2, fluorescent polystyrene, and carbon black nanoparticles on several aquatic crustaceans suggest that nanoparticles are ingested and accumulated in the gut. The particles also adhere to the exoskeleton surfaces of exposed organisms, suggesting multiple routes of exposure can occur simultaneously. CNT induces dose-dependent growth inhibition in a protozoan and is found to be a respiratory toxicant in rainbow trout. Carbon-based nanomaterials are also found to cause toxic effects in fish, including oxidative stress and lipid peroxidation in liver, gill, and brain. Fullerenes and metal (oxide) nanoparticles are also reported to cause delayed hatch, reduced growth, increased molting, oxidative stress, and mortality in various aquatic invertebrates. Current literatures suggest that most sensitive test organism toward nanoparticles are algae and crustaceans and reveal the vulnerability of these organism groups in the aquatic food chain. Besides lethal toxicity, sublethal effects, physiological effects, food web uptake, and interactions with other chemicals should also be investigated.

Toxic effects of various nanomaterials on microorganisms are increasingly being characterized not only because of their great ecological relevance but also due to their attributes as convenient model test organisms. Silver, TiO_2, and ZnO are among the best-studied nanoparticles with respect to microbial toxicity due to their established use as antimicrobial agents. The bactericidal effects of silver compounds and silver ions are well known and have been applied in a wide range of disinfection applications. Release of ionic silver is proposed as one of the major toxic mechanisms of silver nanoparticles although particle effect may also play a role. However, there have been very few comprehensive studies of nanomaterial impacts on environmentally relevant microbial communities. Metal oxide nanoparticles such as TiO_2, SiO_2, and ZnO have toxic effects on bacteria, and the presence of light seems to increase the toxicity.

Information on toxic effects of nanomaterials on marine and estuarine species is scarce. Differences in particle behavior in salt water compared to freshwater impact the fate and transport of nanoparticles and therefore their bioavailability and toxicity. Particle precipitation into sediments under increased salinity results in a lower concentration of smaller particles in the water column and elevated concentrations in sediments. Such different particle behavior and fate between marine and freshwater environments implies that toxicity tests in freshwater are unlikely to provide adequate information on toxicity in seawater and a separate series of tests with marine species may be needed.

It should be noted that results of studies on nanoparticle toxicity depend largely on test conditions. Because such tests have not yet been standardized, it is difficult to make comparisons among species exposed to the same nanomaterials or within the same species using different particle types. A series of important questions need to be answered to pursue standardization of test protocols, such as the following: how do pH, hardness, ionic strength, or organic ligands affect toxicity; is water the correct exposure route or is diet more appropriate; how should particles be dispersed (sonication, stirring, functionalization); and what methods are available for measuring exposure concentration?

Terrestrial Species Much less information is available regarding nanomaterial toxicity to terrestrial organisms compared to aquatic organisms, due to the difficulties in detection of manufactured nanoparticles in soil systems and challenges in obtaining a homogeneous exposure matrix. Terrestrial animals are exposed through dermal and dietary exposure routes, while plants are most likely to be exposed via root uptake.

Toxicological studies on terrestrial species have focused on several metal oxide nanoparticles. For instance, ZnO nanoparticles are found to cause toxicity in the earthworms, and different exposure routes (i.e., dietary vs. dermal) lead to different metal distribution within the animals. Furthermore, additions of humic acids dramatically changed the solubility and morphology and consequently bioavailability and toxicity of the ZnO nanoparticles. Nanosilver toxicity in earthworms is primarily related to the release of Ag ions into the soil solution, and soil properties affect the uptake of Ag from both ionic and particle sources. Uncoated TiO_2 nanoparticles are reported to affect earthworm reproductive activity by abolishing the circannual rhythm that depresses reproduction in the winter, whereas coated TiO_2 does not influence earthworm reproduction. Toxicity of metal oxides, such as TiO_2, ZnO, Al_2O_3, has also been reported in the soil nematode *Caenorhabditis elegans*. ZnO is usually the most toxic among those tested and dissolution to ionic zinc is the main mode of action. Phototoxicity of ZnO nanoparticles to *C. elegans* under natural sunlight has also been reported. More studies into the factors affecting the bioavailability

and toxicity of nanoparticles, as well as the interaction of nanoparticles with soil components, are needed to better understand the potential impacts of these nanoparticles to the soil environments.

Possible interactions of nanoparticles with plant roots are adsorption onto the root surface, incorporation into the cell wall, and uptake into the cell. Exposure to metal nanoparticles has been reported to reduce root elongation and decrease seed germination, depending upon species, test conditions, and concentrations used. In contrast, increased seed germination and photosynthesis in spinach exposed to titanium dioxide nanoparticles have also been reported. CNTs receive attention because of their fibrous structure raising concerns that they may have effects similar to those of asbestos. Observations of the interactions of CNTs with animals, both *in vitro* and *in vivo*, are sometimes contradictory. Similarly, studies assessing CNT toxicity and bioavailability in plants have reported positive, negative, or neutral effects on germination and growth, depending on the test species and CNT surface functionalization. For example, alfalfa and wheat are not damaged by multiwalled CNTs, but rather, their catalytic impurities (Fe species and Al_2O_3) may enhance growth of the plants. These CNTs adsorb onto these plant root surfaces without significant uptake or translocation and do not alter root development or root-tissue morphology. However, evidence on the uptake, accumulation, and generational transmission of NOM-suspended carbon nanoparticles in rice plants suggests the potential impact of nanomaterial exposure on plant development and the food chain and prompts further investigation into the potential impacts of these processes on both food safety and the environment.

Trophic Transfer An important concern in environmental nanotoxicology is the possible propagation of intact nanoparticles and their associated ecological effects through food chains. Bioaccumulation of nanoparticles may occur when nanoparticles from aqueous media enter into organisms, and trophic transfer occurs when the bioaccumulated nanoparticles are passed onto other organisms in a higher trophic level through predation. Biomagnification occurs when the exchange during trophic transfer results in a concentration increase. Evidence has been available for trophic transfer of certain nanomaterials, including QDs, TiO_2, Cu, and Au, among bacteria, algae, protozoans, rotifers, and fish in aquatic environments. Transfer of TiO_2 nanoparticles from daphnids to zebrafish has been reported. Such trophic transfer through terrestrial food webs is also possible. Gold nanoparticles can be transferred from tobacco leaves to tobacco hornworms and biomagnification occurs. As with the legacy contaminants such as methylmercury, DDT, and PCBs, from which dietary uptake at lower trophic levels and accumulation up the food chain represent an important route of exposure to a variety of ecoreceptors as well as humans, trophic transfer and potential biomagnification of nanoparticles in aquatic and terrestrial environments highlight the importance of considering dietary uptake as a pathway for nanoparticle exposure and their potential health risks to ecological receptors and human.

Physicochemical Properties That Affects Toxicity Physical and chemical factors that influence biological effects and pathogenesis of nanomaterials include shape, size, agglomerate/aggregate status, crystallinity, chemical composition, and surface coating. Based on the limited toxicity database of nanomaterials, it appears that surface reactivity is one of the critical factors determining toxicity. As particle size decreases, the surface-to-volume ratio and so the number of reactive groups on the particle surface increase, which are usually responsible for interaction with biological systems and causing toxic effects. Numerous studies in mammalian models have concluded that nanoscale or ultrafine particles (generally defined as particles in the size range < 100 nm) induce greater inflammatory and cytotoxic effects compared to larger-sized particles at equivalent mass concentrations. While the nanosize (and thus increased surface area) alone may contribute to increased surface reactivity for some nanomaterials, other physicochemical properties (e.g., dimension and shape, agglomeration state, chemical composition, crystal structure, surface coating) may be modifiers and important determinants to their surface reactivity and thus their toxicological potential. For example, pulmonary exposure to nanoscale TiO_2 particles does not produce more cytotoxic or inflammatory effects to the lungs of rats compared with fine-sized TiO_2, indicating that toxicity is not always dependent upon particle size and surface area.

Nanomaterials have various shapes such as spherical, needlelike, tubes, rods, etc. The shape of nanomaterials may have effects on their kinetics of deposition and absorption to biological systems and thus affect their toxicological effects. Furthermore, most nanoparticles tend to form aggregates or agglomerates in environmental and biological media, and the structures/states of these aggregates/agglomerates may affect toxicity. Aggregation of ZnO nanoparticles in aquatic media significantly decreases their toxicity to zebrafish embryo/larvae due to decreased bioavailability. Different aggregate/agglomerate structures of single-walled CNTs have been associated with distinct regional responses of mice lungs because of their different size, shape, and deposition/clearance kinetics. Surface modification of nanoparticles can significantly alter their physicochemical properties and consequently modify their biological effects. For instance, hydroxylation of fullerene can render the nanomaterial much less toxic than its pristine form and induce distinct types of cell death through different mechanisms. Coating of QDs with different types and concentrations of surfactants has also been shown to change their distribution and effects on biological systems significantly.

Although most nanomaterials are poorly soluble, dissolution of some nanoparticles occurs in culture medium or biological fluids. Such dissolution can have significant impact on the cellular uptake, subcellular localization, and toxic effects of these nanoparticles. This is especially the case for metal or metal oxide nanoparticles such as silver, ZnO, and CuO. Silver's antimicrobial activity is most often attributed to the dissolved ions rather than the high surface area, low-solubility nonionic metallic nanoparticles.

19.5 MECHANISMS OF TOXICITY (MODES OF ACTION)

While toxicity mechanisms have not yet been completely elucidated for most nanomaterials, possible modes of action include the production of intracellular ROS and related oxidative stress, disruption of membranes or membrane potential, oxidation of proteins, genotoxicity, interruption of energy transduction, and release of toxic constituents. The nano–bio interface is a dynamic system where many short- and long-term interactions occur simultaneously as well as sequentially. Figure 19.2 depicts the possible mechanisms by which nanomaterials interact with a biological system. The examples illustrated demonstrate the importance of material composition, electronic structure, surface chemistry, solubility, as well as environmental factors (e.g., UV radiation) in determining nano–bio interaction and subsequent toxicological effects.

Oxidative Stress

Generation of ROS and subsequent oxidative stress is currently the best-developed paradigm for nanoparticle toxicity, as several nanomaterial characteristics (e.g., small size, increased number of structural defects, altered electronic properties) facilitate in ROS generation. The capability of nanomaterials to generate ROS is greatly associated with the change in physicochemical and structural properties of materials in nanoscale that are responsible for their interaction with environmental factors and biological

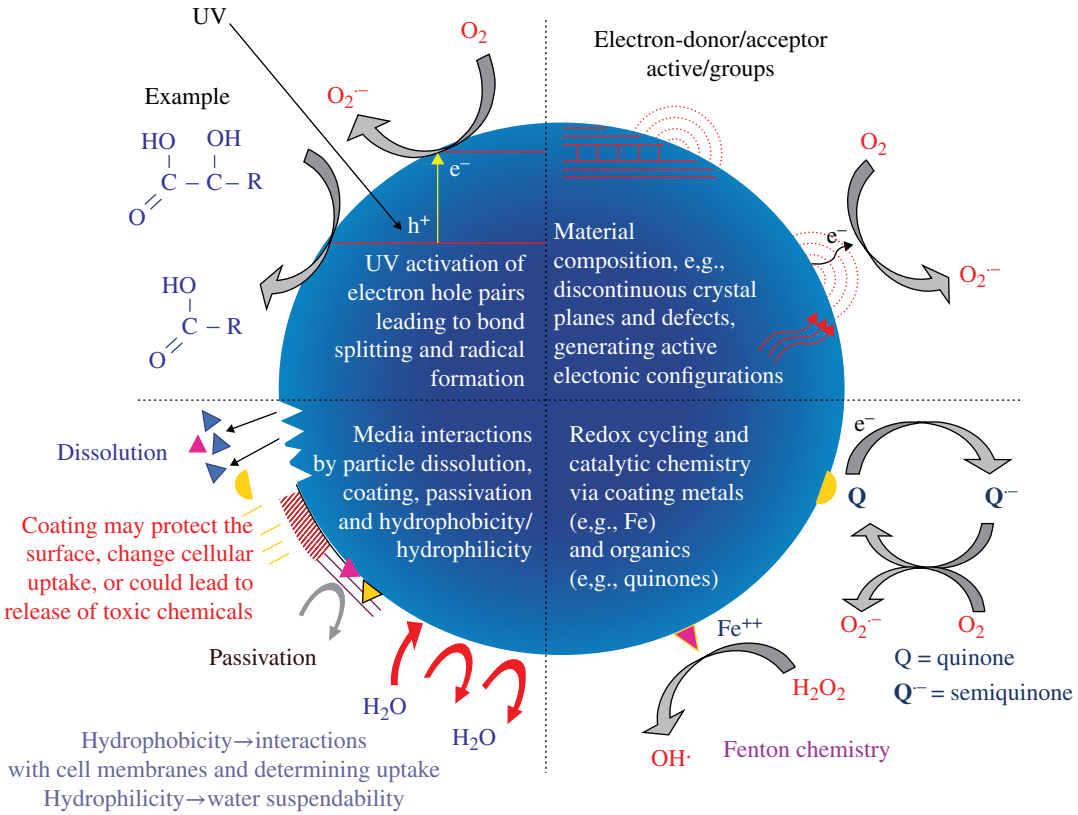


FIGURE 19.2 Possible mechanisms by which nanomaterials interact with biological tissue. Examples illustrate the importance of material composition, electronic structure, bonded surface species (e.g., metal containing), surface coatings (active or passive), and solubility, including the contribution of surface species and coatings and interactions with other environmental factors (e.g., UV activation). *Source:* Reprinted with permission from Nel et al. (2006). © 2006 AAAS.

systems. For instance, shrinkage in size can increase the number of structural defects and alter electronic properties of the material, which could establish specific surface groups that may function as reactive sites for ROS generation. Such electron donor or acceptor active sites (chemically or physically activated) interacting with molecular oxygen will generate superoxide radical through electron capture (Figure 19.2). The presence of transition metals on the surface of single-component materials can participate in the formation of such active sites, such as ultrafine particles containing transition metals or coated with redox-cycling organic chemicals (Figure 19.2).Oxidative stress occurs when a greater number of oxidants than antioxidants are present within the cell, causing an oxidant/antioxidant imbalance. Low level of oxidative stress is associated with the induction of antioxidant and detoxification enzymes. As the oxidative stress reaches at higher levels, this protective response is overtaken by inflammation and cytotoxicity. A number of responses at each level of oxidative stress have now been successfully incorporated as screening assays for toxicological effects of ambient particulate matter *in vivo*. Characterization of particle size and physical properties, together with *in vitro* assay for ROS and oxidative stress plus *in vivo* markers of oxidative stress, is an example of predictive paradigm for toxicity screening.

Nanoparticles of various chemical compositions have been shown to generate ROS in both *in vivo* and *in vitro* studies. Nanoparticles have also been implicated in interfering with cell signaling via ROS-mediated activation of cytokine gene expression. In mammalian studies, there is a direct relationship between the surface area, ROS-generating capability, and proinflammatory effects of nanoparticles in the lung. Antibacterial activity of nanoparticles occurs through a direct contact between nanoparticles and the bacterial cell surface, which affects the integrity of cell membrane and causes oxidative stress and formation of highly reactive epoxides. These epoxides can induce DNA damage within the nucleus.

Dissolution to Toxic Ions

Solubility strongly influences toxicity of metal and metal oxide nanoparticles. The bactericidal effect of silver nanoparticles is due to the release of silver ions. ZnO is usually found to be most toxic among other metal oxide nanoparticles (e.g., CuO, TiO_2, SiO_2) due to its high solubility and the high potency of dissolved zinc ions to aquatic organisms. The impact of solubilization from metal-containing nanoparticles can be studied using metal-specific microbial sensor. Those genetically modified microbial biosensor strains only produce a response if the toxicant crosses the cell biological envelopes and enters the cytoplasmic space. This approach has been successfully used for the demonstration of toxic effects of solubilized Zn and Cu from nanoparticles of ZnO and CuO toward algae, crustaceans, and bacteria, protozoa, and yeast.

Particle dissolution is dependent on both the intrinsic physicochemical properties of the material (e.g., particle size, surface area, chemical composition) and the environmental parameters of the exposure matrix (e.g., pH, temperature, organic matter). Particles with smaller size are expected to have increased dissolution as a result of their increased surface area. pH also has a significant impact on particle dissolution with lower pH (more acidic conditions) favoring dissolution. Influence of NOM on particle dissolution is usually dependent on specific chemical structure of the NOM. For example, NOM may enhance ZnO dissolution by providing chelating agents for zinc ions or inhibit dissolution by adsorption to the particle surface and blocking it from interaction with water molecules. Therefore, toxicity test results for relative soluble metal-containing nanoparticles are highly dependent on test conditions. Water solubility of nanoparticles has to be incorporated into the environmental and human health risk assessment models of nanoparticles in addition to other key physicochemical characteristics relevant to nanoparticles.

Others

In addition to the paradigm of oxidative stress and particle dissolution to toxic ions, nanomaterials with different physicochemical properties can cause different types of inflammation and toxic responses by different mechanisms. These may include protein denaturation, membrane damage, DNA damage, immune reactivity, and the formation of foreign body granuloma. Nanoparticles with large surface areas may activate receptors and trigger various signal transduction pathways after interaction with plasma membrane. They can also enter cells via endocytosis and interfere with normal cellular functions through interaction with intracellular macromolecules. Potential nanoparticle–cellular interactions that may induce cytotoxicity and other cellular responses include (i) interaction with plasma membrane, which may cause instability associated with ion transport, signal transduction, and cell death; (ii) interaction with mitochondria, which may alter metabolism or interfere with antioxidant defenses and ROS production; (iii) binding to DNA, which may damage DNA, cell cycle division, and protein synthesis; (iv) interaction with cytoskeleton, which may halt vesicular trafficking and cause mechanical instability and cell death; and (v) interaction with proteins, lipids, and other biomolecules, which may lead to different types of "corona" and biological effects. A summary of possible mechanism of toxicity for nanomaterials as suggested from experimental and clinical evidence is shown in Table 19.2.

TABLE 19.2 Possible Mechanisms of Toxicity for Nanomaterials (Including Both Naturally Occurring and Man-Made) as Suggested from Experimental and Clinical Evidence

Experimental Nanomaterial Effects	Possible Pathophysiological Outcomes
ROS generation*	Protein, DNA, and membrane injury,* oxidative stress†
Oxidative stress*	Phase II enzyme induction, inflammation,† mitochondrial perturbation*
Mitochondrial perturbation*	Inner membrane damage,* permeability transition (PT), pore opening,* energy failure,* apoptosis,*aponecrosis, cytotoxicity
Inflammation*	Tissue infiltration with inflammatory cells,† fibrosis,† granulomas,† atherogenesis,† acute phase protein expression (e.g., C-reactive protein)
Uptake by reticuloendothelial system*	Asymptomatic sequestration and storage in liver,* spleen, lymph nodes,† possible organ enlargement and dysfunction
Protein denaturation, degradation*	Loss of enzyme activity,* autoantigenicity
Nuclear uptake*	DNA damage, nucleoprotein clumping,* autoantigens
Uptake in neuronal tissue*	Brain and peripheral nervous system injury
Perturbation of phagocytic function,* "particle overload," mediator release*	Chronic inflammation,† fibrosis,† granulomas,† interference in clearance of infectious agents†
Endothelial dysfunction, effects on blood clotting*	Atherogenesis,* thrombosis,* stroke, myocardial infarction
Generation of neoantigens, breakdown in immune tolerance	Autoimmunity, adjuvant effects
Altered cell cycle regulation	Proliferation, cell cycle arrest, senescence
DNA damage	Mutagenesis, metaplasia, carcinogenesis

Source: Reprinted with permission from Nel et al. (2006). © 2006 AAAS.
Effects supported by limited experimental evidence are marked with asterisks; effects supported by limited clinical evidence are marked with daggers.

19.6 CHALLENGES AND FUTURE DIRECTIONS

Nanotoxicology as a new discipline is anticipated to make an important contribution to the development of a sustainable and safe nanotechnology. Toxicological studies are the basis for protection of human health and the environment relating to nanotechnology. It is only through addressing the issues raised by toxicological studies that nanotechnology will be able to realize its full potential. There has been significant progress in the field of nanotoxicology since the production of nanotechnology-based products. However, the field has not been able to maintain an even balance in regard to the known effects of nanoparticles and the nanotechnology-based productions and applications. Nanotoxicology studies require a standard set of protocols for *in vitro* and *in vivo* toxicity and ecotoxicity. Quantitative data on toxicological effects of nanoparticles are still scarce even at the single organism level and are far less sufficient for risk assessment and regulatory purposes. Currently, neither the fate of nanosized materials nor their impact on animals, plants, and soil communities has been investigated *in situ* although it would be necessary for the validation of models derived from laboratory studies that are proposed for risk assessment of nanomaterials.

Additional studies are needed to clarify whether existing methodologies for assessing ecological and mammalian toxicological effects are adequate for nanomaterials or if alternatives need to be developed. It is widely accepted that the traditional methods of measuring dose exposure are of little use for predicting the toxicological effects of nanomaterials. Standard monitoring instruments cannot always detect nanoparticles in environmental and biological samples. Thus, characterizing their behavior and novel properties, and tracing their impacts, presents a real scientific and technical challenge.

In addition to standard tests, there is a need to develop better and rapid screening methods and to move into more predictive toxicology. Before the manifestation of acute toxic effects on the cellular or organismal level, toxicant-specific changes occurring at the molecular level may provide information about toxicity pathways. A significant parallel between "omics" data and effects endpoints suggests that "omics" may help to elucidate potential mechanisms of toxicity. The desire to reduce, refine, and replace animal use in toxicological research creates an urgent need to develop mechanism-centered high-throughput testing procedures. As alternative methods become more accessible and widely used, functional toxicogenomics will play an important role in identifying the essential cellular components and pathways involved in toxicity responses to nanomaterials.

19.7 SUMMARY

The increasing use and production of nanomaterials lead to multiple potential points of entry of the nanomaterials to the environment, resulting in human and ecosystem exposure. These include the exposure from workplace (manufacturing), during product use, and during waste cycle.

Several studies have demonstrated a release of engineered nanomaterials incorporated in textiles and paints to the aquatic environments, with a significant fraction of particles being found to escape the clearing system of wastewater plants. Due to the expanding use of nanomaterials in the food industry, personal care products, and the field of medicine, human exposure to nanomaterials on a daily basis is a near certainty. It is important that we understand the ecological and human health impacts of new nanomaterial uses for regulatory and risk assessment purposes.

The number of toxicological studies of engineered nanomaterials has rapidly increased in the past few years. However, the data obtained so far are somewhat inconsistent and not sufficiently systematic to allow an overview of the potential environmental and human health hazards relating to engineered nanomaterials. In addition to an overall lack of well-rounded studies, inadequacies in material characterization for many publications have limited their contribution to our understanding of this topic. In order to create a meaningful database of nanotoxicity studies, it is critical for researchers to consider appropriate characterization, dosimetry, and endpoint measurement when designing an experiment.

SUGGESTED READING

Adams LK, Lyon DY, Alvarez PJJ. Comparative eco-toxicity of nanoscale TiO_2, SiO_2, and ZnO water suspensions. Water Res 2006;40:3527–3532.

Arora S, Rajwade JM, Paknikar KM. Nanotoxicology and *in vitro* studies: the need of the hour. Toxicol Appl Pharmacol 2012; 258(2):151–165.

Baer DR, Gaspar DJ, Nachimuthu P, Techane SD, Castner DG. Application of surface chemical analysis tools for characterization of nanoparticles. Anal Bioanal Chem 2010;396(3):983–1002.

Claverie C, Corbella R, Martín D, Díaz C. Protective effects of zinc on cadmium toxicity in rodents. Biol Trace Elem Res 2000;75(1): 1–9.

Gottschalk F, Sonderer T, Scholz RW, Nowack B. Modeled environmental concentrations of engineered nanomaterials (TiO_2, ZnO, Ag, CNT, Fullerenes) for different regions. Environ Sci Technol 2009;43(24):9216–9222.

Isakovic A, Markovic Z, Todorovic-Markovic B, Nikolic N, Vranjes-Djuric S, Mirkovic M, Dramicanin M, Harhaji L, Raicevic N, Nikolic Z, Trajkovic V. Distinct cytotoxic mechanisms of pristine versus hydroxylated fullerene. Toxicol Sci 2006;91(1):173–183.

Jiang J, Oberdorster G, Biswas P. Characterization of size, surface charge, and agglomeration state of nanoparticle dispersions for toxicological studies. JNR 2009;11:77–89.

Judy JD, Unrine JM, Bertsch PM. Evidence for biomagnification of gold nanoparticles within a terrestrial food chain. Environ Sci Technol 2011;45:776–781.

Kahru A, Dubourguier HC. From ecotoxicity to nanoecotoxicology. Toxicology 2010;269:105–119.

Klaine SJ, Alvarez PJJ, Batley GE, Fernandes TF, Handy RD, Lyon DY, Mahendra S, McLaughlin MJ, Lead JR. Nanomaterials in the environment: behavior, fate, bioavailability, and effects. Environ Toxicol Chem 2008;27(9):1825–1851.

Li LZ, Zhou DM, Peijnenburg WJM, van Gestel CAM, Jin SY, Wang YJ, Wang P. Toxicity of zinc oxide nanoparticles in the earthworm, *Eisenia fetida* and subcellular fractionation of Zn. Environ Int 2011;37(6):1098–1104.

Lin DH, Xing BS. Phytotoxicity of nanoparticles: inhibition of seed germination and root growth. Environ Pollut 2007;150(2): 243–250.

Ma H, Bertsch PM, Glenn TC, Kabengi NJ, Williams PL. Toxicity of manufactured zinc oxide nanoparticles in the nematode *Caenorhabditis elegans*. Enviorn Toxicol Chem 2009;28 (6):1324–1330.

Ma H, Kabengi NJ, Bertsch PM, Unrine JM, Glenn TC, Williams PL. Comparative phototoxicity of nanoparticulate and bulk ZnO to a free-living nematode *Caenorhabditis elegans*: the importance of illumination mode and primary particle size. Environ Pollut 2009;159(6):1473–1480.

Masciangiol T, Zhang W. Environmental technologies at the nanoscale. Environ Sci Technol 2003;37(5):102A–108A.

Maynard AD. *Nanotechnology: A Research Strategy for Addressing Risk. Project on Emerging Nanotechnologies*. Washington (DC): Woodrow Wilson International Center for Scholars; 2006.

Navarro E, Baun A, Behra R, Hartmann N, Filser J, Miao AJ, Quigg A, Santschi P, Sigg L. Environmental behavior and ecotoxicity of engineered nanoparticles to algae, plants, and fungi. Ecotoxicology 2008;17(5):372–386.

Nel A, Xia T, Madler L, Li N. Toxic potential of materials at the nanolevel. Science 2006;311(5761):622–627.

Nohynek GJ, Lademann J, Ribaud C, Roberts MS. Grey goo on the skin? Nanotechnology, cosmetic and sunscreen safety. Crit Rev Toxicol 2007;37(3):251–277.

Nowack B, Bucheli TD. Occurrence, behavior and effects of nanoparticles in the environment. Environ Pollut 2007;150: 5–22.

Oberdorster G, Oberdorster E, Oberdorster J. Nanotoxicology: an emerging discipline evolving from studies of ultrafine particles. Environ Health Perspect 2005;113:823–839.

OECD Environment, Health and Safety Publications. 2005–2014. *Series on the Safety of Manufactured Nanomaterials*. Available at www.oecd.org/ehs/. Accessed August 5, 2014.

Shoults-Wilson WA, Reinsch BC, Tsyusko OV, Bertsch PM, Lowry GV, Unrine JM. Toxicity of silver nanoparticles to the earthworm (*Eisenia fetida*): the role of particle size and soil type. Soil Sci Soc Am J 2011;75:365–377.

Stoeger T, Reinhard C, Takenaka S, Schroeppel A, Karg E, Ritter B, Heyder J, Schulz H. Instillation of six different ultrafine carbon particles indicates a surface area threshold dose for acute lung inflammation in mice. Environ Health Perspect 2006;114(3): 328–333.

Stone V, Johnston H, Schins RRF. Development of *in vitro* systems for nanotoxicology: methodological considerations. Crit Rev Toxicol 2009;39(7):613–626.

Werlin R, Priester JH, Mielke RE, Kramer S, Jackson S, Stoimenov PK, Stucky GD, Cherr GN, Orias E, Holden PA. Biomagnification of cadmium selenide quantum dots in a simple experimental microbial food chain. Nat Nanotechnol 2011;6(1):65–71.

Wiesner MR, Lowry GV, Alvarez P, Dionysiou D, Biswas P. Assessing the risks of manufactured nanomaterials. Environ Sci Technol 2006;40(14):4336–4345.

Wittmaack K. In search of the most relevant parameter for quantifying lung inflammatory response to nanoparticle exposure: particle number, surface area, or what? Environ Health Perspect 2007;115(2):187–194.

20

COMPUTATIONAL TOXICOLOGY*

RICHARD S. JUDSON, DAVID M. REIF, AND KEITH A. HOUCK

This chapter provides an overview of the emerging field of computational toxicology (CT) and discusses:

- Goals of CT research and development
- Types of data generated and used
- Organization of CT data into databases and knowledge bases
- Applications

CT is an emerging field that combines *in vitro* and computationally generated data on chemicals, information on biological targets (genes, proteins), pathways and processes, and informatics methods to model and understand the mechanistic basis of chemical toxicity. CT often looks at trends across large sets of chemicals, large sets of data on a single chemical, or a combination of the two. Much of the computational effort focuses on organizing these data and using statistical and modeling methods to interpret them. Two other areas of research that often fall under the CT heading, but which will not be covered here, are systems biology modeling and quantitative structure–activity relationship (QSAR) modeling. These purely computational approaches are complementary to the *in vitro*, data-centered approach described here but are sufficiently different to warrant their own in-depth discussion. This chapter will introduce the goals, tools, and approaches used by CT practitioners and will illustrate them through several examples.

20.1 DATA RELEVANT TO CT APPLICATIONS

Although the term CT may imply a purely theoretical approach, in reality, all CT methods are heavily reliant on data to drive hypotheses and to build and validate models. What distinguishes many of the experimental aspects of CT from those of traditional toxicology is the reliance on high-throughput, *in vitro* methods. These methods are seeing increases in both usage and reliability. In addition to producing new data, CT methods often require animal-based *in vivo* data to anchor predictions from *in vitro*-based models. To do this in an efficient manner, it is important to compile quantitative data from a large number of animal studies into "computable" databases. Here, computable means that numerical or categorical data are extracted from text reports and tabulated.

Chemical Identity, Structure, and Properties

Like most toxicology methods, CT focuses on the potential toxic effects of chemicals and attempts to build models linking chemicals and their structures and properties to *in vitro* activity and *in vivo* toxicity. To do this, chemicals under study must be accurately identified by names or other identifiers and by chemical structure. Although seemingly straightforward, arriving at unique and correct chemical identity requires significant care.

Many chemicals go by common names but can also be identified by systematic names [i.e., using IUPAC rules

*__Disclaimer__: *The views expressed in this paper are those of the authors and do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency.*

Principles of Toxicology: Environmental and Industrial Applications, Third Edition. Edited by Stephen M. Roberts, Robert C. James and Phillip L. Williams.

(IUPAC, 1993)] and a variety of trade names. For instance, the pesticide atrazine has over 100 synonyms listed in public databases, including more than 25 variant systematic names. Some of this ambiguity reflects product names from different manufacturers and different formulations with greater or lesser degrees of purity. The Chemical Abstracts Service Registry Number (CAS registry number or CASRN) is a widely used alternative to uniquely identify a chemical, but even here, there is ambiguity. For atrazine, the accepted CASRN is 1912-24-9, but a variety of sources list at least six others. Some of these are old (withdrawn) CASRN, some refer to mixtures, and some are simple mistakes that have crept into papers and online databases. The official database of CASRN is managed by Chemical Abstracts Service and requires fee-based access. The European Commission, through the Joint Research Centre, has developed an alternative identifier called the EC number, which for atrazine is 217-617-8. This is an openly available system, but unfortunately, few data in the toxicology literature are annotated with EC numbers. Despite these drawbacks, several publicly available databases have carefully curated CASRN, systematic names, and structures for many chemicals. These include DSSTox (Richard and Williams, 2002; Russom et al., 2008) and ChemSpider (2008).

Pure chemicals can also be uniquely identified by their structure. There are three common structure conventions, plus a plethora of others. These are the Simple Molecular Input Line Entry System (SMILES) (Daylight, 2008), IUPAC International Chemical Identifier (InChI) (IUPAC, 2008), and Mol or structure definition (SD) file. SMILES and InChI are both codes that can be written on a single line and can be used to reconstruct the two-dimensional geometry of a chemical. For atrazine (Figure 20.1), the SMILES and InChI codes are "CCNc1nc(nc(n1)Cl)NC(C)C" and "InChI=1S/C8H14ClN5/c1-4-10-7-12-6(9)13-8(14-7)11-5(2)3/h5H,4H2,1-3H3,(H2,10,11,12,13,14)," respectively. Because the InChI codes can be very long, an alternative short version called the InChIKey has been developed. This is a "one-way" descriptor, meaning that one can generate the key from a structure, but cannot go backward. Therefore, it is useful as a unique identifier, but not for use as a shorthand to reconstruct structures for later analyses. For atrazine, the InChIKey is MXWJVTOOROXGIU-UHFFFAOYSA-N. SMILES are currently more widely used than InChIs, but in the future, this may change because the InChI code has the advantage of being unique—that is, the description of the coding algorithm enforces uniqueness. With SMILES, on the other hand, one can generate multiple valid strings to represent the same molecule. There is ongoing work in both the InChI and SMILES community to develop more robust coding algorithms that guarantee (for instance) the correct handling of structures with multiple chiral centers. There have been proposals to replace the use of arbitrary identifiers such as CASRN with the InChIKeys, but this begs the question of how to deal with mixtures or formulations, which are common subjects of chemical and toxicological study.

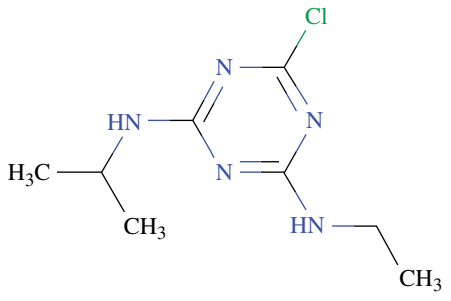


FIGURE 20.1 Chemical structure of atrazine.

Given a chemical structure, one can calculate or look up many properties that are often important for toxicology studies. These include the octanol–water partition coefficient (logP or logKow), water solubility, melting and boiling point, etc. A commonly used, freely available tool is EPI Suite (U.S. EPA, 2010), which has the advantage of containing a large database of experimental values and will give the user an experimentally derived value when available rather than just returning a calculated estimate. EPI Suite is especially rich in data on environmental chemicals such as pesticides and industrial chemicals. It also provides estimates of several parameters used in ecotoxicology studies. A variety of commercial property-estimating packages are also widely used, including MOE (http://www.chemcomp.com/software.htm), QikProp (Schrodinger, 2010), Leadscope (2010), and OpenEye Babel (http://www.eyesopen.com/docs/babel/current/html/index.html).

In Vitro Methods and Data

In vitro assays are a key source for data for many CT applications. These assays test chemicals using cell-based or cell-free assay systems. This approach allows one to understand the effects of chemicals at a molecular and cellular level, including direct interactions with biomolecules (DNA, proteins) and the molecular or pathway-based effects triggered by chemical exposures to cells. Here, we describe the most widely used *in vitro* techniques for CT, which are high-throughput screening (HTS) and high-content screening (HCS) and whole-genome analyses. Other techniques not covered here are proteomics, in which one measures broad-based protein level changes in response to a chemical, and metabolomics, where the same type of experiment is performed, only looking at natural metabolites or small molecules.

Cell Systems (Primary, Cell Lines, Tissues, Cocultures, and Cell-Free)

The choice of cell system to use in any cell-based assay is critical. This is because of the tissue-specific expression of genes and proteins, activity of signaling pathways, presence

of necessary cofactors, etc., that determine whether the target of toxicity exists in the selected cells. Choices of cell systems include primary cells, cell lines, stem cells, cocultures (i.e., mixtures of different cell types), tissue slices or sections (usually *ex vivo*), and three-dimensional versions of some of these.

Primary cells are isolated directly from an animal and have the advantage of behaving in a physiologically normal way, at least for a brief period. Signaling pathways are intact, there are a normal number of chromosomes, and xenobiotic metabolism can be active. There are, however, several drawbacks to using primary cells, the most obvious of which is the difficulty of obtaining these cells, especially for human tissues outside of blood cells, combined with the often limited ability to passage these cells in culture. Additionally, normal physiological function typically decreases in 24–48 h, so that any experiments need to be completed in this short window of time. Further, there can be significant donor-to-donor variability, even when cells come from inbred animals. The donor-variability issue is even larger with human samples.

Immortalized cell lines address some of the issues with primary cells, mainly that they can be much less variable because a large volume of nearly identical cells can be obtained that can be used across time and across labs. The major downside to using cell lines is that they are by definition abnormal, being most often derived from cancer cells and adapted to growth *in vitro* with resulting major changes in important signaling pathways controlling growth and differentiation. They are usually significantly different from normal cells derived from the same tissue as exemplified by the greatly reduced xenobiotic enzyme content and inducibility in the HepG2 human liver cell line relative to normal hepatocytes (Westerink and Schoonen, 2007).

A great advantage of cell lines is the ability to compare results lab to lab and year to year, but it is known that there is genetic (and probably epigenetic) drift in cell lines over passage number. While there are literally hundreds of cell lines derived from numerous tissues available, offering many choices for models, the lack of standardization often challenges the ability to compare results systematically. However, the diversity of cell lines available provides interesting opportunities, exemplified by the lines created for the human HapMap project (The International HapMap Consortium, 2005). These are immortalized lymphocytes from an ethnically diverse set of individuals, which have been genotyped for over 500K single nucleotide polymorphisms (SNPs). This resource allows researchers to study the effects of genetic variation on the activity of chemicals on cells (O'Shea et al., 2010).

Pluripotent and multipotent stem cells are becoming more widely used in toxicology because they can have the advantages of being more "normal" than typical cell lines and can be made to differentiate into many other cell types. Today, embryonic stem cells are most often used (either mouse or one of the limited number of human lines that are widely allowed to be used). At the forefront in toxicology applications is the derivation of cardiomyocytes for use in testing for cardiotoxic effects of chemicals (Dick et al., 2010). There is also high interest in developing liver cells from pluripotent stem cells, which would have wide applications in toxicology (Huang et al., 2011). The ability to create induced pluripotent stem cells from adult tissue promises to make more lines widely available for toxicity research in the near future.

There are a number of coculture systems that are being used—these are mixtures of different cell types in either 2D or 3D cultures. The reason for using such cocultures is that they can be used to form complete signaling cascades that are characteristic of normal tissues composed of multiple cell types. These cocultures are sometimes referred to as organotypic. For example, culture systems have been designed to model pathophysiology of inflammation and cardiovascular and respiratory diseases (Berg et al., 2010). Primary tissue analogs of these organotypic cultures are *ex vivo* tissue slices such as the hippocampal slices used to study brain damage (Noraberg et al., 2005) or liver slices for studying hepatotoxicity (Boess et al., 2003). We should also mention in passing that certain model organisms are being used in the CT field, mainly zebrafish embryos and *C. elegans* (Parng et al., 2002; Smith et al., 2009). These systems allow interrogation of higher-level, emergent properties of intact organisms that are amenable to treatment and study in a medium-throughput format in 96-well microtiter plates.

HTS Assays, Technologies, Trends

HTS refers to a set of techniques that measure interactions of chemicals with proteins or cells and does this for many chemicals or conditions simultaneously, typically in microtiter plates. Many of these assays were initially developed by or for the pharmaceutical industry for the testing of thousands to millions of compounds against molecular targets to find lead compounds against specific diseases (Bleicher et al., 2003; Mayr and Bojanic, 2009). Assays are typically run in standardized 96-, 384-, or 1536-well plates, often in concentration–response mode. With increased density of wells per plate, the quantity of cells and reagents and the cost per chemical all decrease, providing vast increases in efficiency. In order to manage these large numbers of samples, automation is often used, ranging from simple liquid handling stations to fully automated robotic systems that can fill an entire room. Each assay is characterized by whether it is run in cell-free, that is, biochemical, format or with cultured cells; by its target; and by its signal readout. Biochemical assays targeting specific protein interaction can measure chemicals binding to receptors, chemicals

FIGURE 20.2 A large-scale robotic screening system in use at the NIH Chemical Genomics Center. This is capable of screening up to 100,000 compounds at a time. Inside the enclosure is a large robot arm that moves plates between stations.

interfering or promoting protein–protein interactions, or chemicals affecting enzymatic activity. Assay signal readouts are typically generated through the use of radiolabeled or fluorescently labeled compounds, although for HTS, radiolabeled tests are less often used due to the volume of radioactive waste generated (Sundberg, 2000). A wide variety of assay technologies are commercially available to measure these activities. A recent trend is toward the use of label-free technologies such as high-throughput mass spectrometry to avoid the inherent changes to proteins and substrates involved with fluorescent and other tags (Lunn, 2010). In addition to biochemical assays, the use of cell-based assays in HTS format has become routine. These cellular assays are even used in ultrahigh-throughput approaches as exemplified by the quantitative HTS (qHTS) approach developed by the NIH Chemical Genomics Center (NCGC). The NCGC is able to rapidly test libraries of hundreds of thousands of compounds at 12–15 concentrations, so that quantitative concentration–response curves are generated (Inglese et al., 2006, 2007). This is of particular importance to the toxicology field where understanding dose-related activity is critical (see Figure 20.2 for an example of the large robotic systems used in HTS testing).

HTS methods are distinguished across several axes. First is what is being measured, most often falling into the following categories:

- *Transcript levels*: Detect changing levels of mRNA transcripts in cells, for instance, in response to a chemical interacting with a nuclear receptor or elsewhere in a transcription factor pathway. The number of different transcripts measured depends on the technology and ranges from one or two in a reporter gene assay to tens, hundreds, or thousands with other methods. Reporter gene assays indirectly measure engineered mRNA transcripts encoding reporter genes such as luciferase or β-lactamase through the activity of the reporter gene product. The other methods measure endogenous mRNA transcripts directly or after conversion to cDNA and amplification. Microarray chip-based hybridization/fluorescence techniques can be used for a few genes or a whole genome but have somewhat limited dynamic range. An alternative is real-time polymerase chain reaction (RT-PCR), which is a more quantitative technique for measuring transcript levels and is typically run for up to 40 transcripts at a time. Chip or bead-based quantitative nuclease protection assays (QNPA) (Roberts et al., 2007; Rimsza et al., 2008) are another improvement over microarray chip-based methods because they do not require a PCR amplification step, which reduces cost and increases quantitative reproducibility. QNPA can be multiplexed with up to 200 transcripts measured in a single well.
- *Protein levels*: The most often used technique is ELISA, or enzyme-linked immunosorbent assay. An antibody is developed for each protein to be detected. This is more difficult and expensive than equivalent RNA detection techniques because of the need to develop highly specific antibodies (Taipa, 2008).
- *Protein activity*: These assays are usually run cell-free and measure the change in activity (usually for an enzyme target) when a test chemical is introduced. In the most common format, the production or degradation of a fluorescent substrate resulting from enzymatic activity of the target protein is monitored and inhibition of that enzymatic activity by the test chemical is determined.
- *Protein binding*: These assays are typically run cell-free and measure the ability of a ligand to bind to a specific protein. Most often, a strong native ligand is initially bound to the protein and one measures how well the test ligand displaces the native ligand, which is usually fluorescently or radiolabeled. Effect of chemicals on protein–protein binding interactions can also be measured by techniques such as fluorescent resonance energy transfer (FRET), which requires fluorescent labeling of both target proteins.
- *Cytotoxicity*: The killing of cells by chemical treatment can be monitored by a variety of techniques including measurement of mitochondrial reductase activity, loss of cell membrane integrity, and loss of ATP content. In any type of cell-based assay, it is important to monitor for cytotoxicity over the same concentration range of the test chemical as is used in the direct assay to control for possible interference with the measured endpoint. Ideally, one measures ligand activity and cytotoxicity in the same well.

TABLE 20.1 Typical Types of Readouts Used in HTS Assays

Readout Type	Description or Use	Advantages	Disadvantages
Radioligand	Typical use is in cell-free assays to measure binding or activity of small molecules against receptors or enzymes. The target protein is preincubated with a radiolabeled ligand. The test chemical is then added and the amount of displaced radioligand is measured as a function of test chemical concentration	Sensitive, directly measures chemical product	Requires synthesis, use, and disposal of radioligand
Fluorescence	Uses include assays that measure enzymatic activity by effects of enzyme on fluorescently labeled substrate, cellular production of fluorescent protein, labeling of cDNA to detect levels of specific mRNAs, and measurement of protein complexes, which then fluoresce due to fluorescent resonant energy transfer between labeled molecules	Highly sensitive detectors, some multiplexing (more than one color per well), inexpensive	Interference by fluorescent test chemicals (fluorescent intensity and fluorescent quenching) generating potential false-positive and false-negative results, background cellular fluorescence
Luminescence	Luciferase reporter gene product signal, enzyme substrates (become luciferase substrate after enzyme activity)	Very low background, high dynamic range	Test chemicals that inhibit luciferase enzymatic activity
Light absorption	Chloramphenicol acetyltransferase reporter gene product signal, enzyme substrates (become colored after enzyme activity)	Inexpensive and simple	Often poor sensitivity and limited dynamic range, test chemical interference with absorption at same wavelength
Nucleic acid sequence	Detects changing levels of mRNA transcripts, for instance, in response to a chemical interacting with a nuclear receptor or elsewhere in a transcription factor pathway	Can be highly multiplexed, sensitive	Can be expensive; some methods have limited sensitivity and dynamic range
Antibody protein	Detects the level of a target protein (ELISA)	Can be multiplexed, directly detects protein levels	Requires production of a specific antibody; sensitivity often limiting; assay usually requires wash steps
Cellular impedance	Change in cell number or shape affecting impedance in a microelectrode array incorporated into a microtiter plate	Sensitivity, no label required, real-time measurement	Limited endpoints detected

The next axis is the type of system (cell-based or cell-free):

- *Biochemical (cell-free) or cellular*: In a biochemical system, one is looking for interactions of a chemical with a protein (e.g., a receptor or enzyme). In cell-based systems, one can still measure direct chemical–protein interactions but can also monitor molecular or cellular changes of chemical exposure.
- *Cell type*: The types of cells used can greatly affect the types of results one can measure.

Finally, one must consider the readout used in the assay. The effect of a chemical exposure can be monitored using levels of florescence, luminescence, light absorption, radiolabel methods, cell impedance, cell imaging (covered in Section "High-content methods"), antibody-based protein detection (ELISA), or sequence readouts. Each of these has advantages and disadvantages, annotated in Table 20.1.

High-Content Methods

HCS methods make use of automated fluorescent microscopy techniques (Bullen, 2008; Giuliano et al., 2005, 2006) to obtain information on cell size, morphology, or subcellular localization of proteins or other biomolecules. One typically fluorescently stains cells using dyes with unique excitation/emission spectra that target subcellular organelles (nucleus, cytoplasm, cell membrane, etc.) and simultaneously labels specific target proteins using antibodies

conjugated with fluorescent tags of distinct excitation/emission spectra. The stained cells are imaged on the fluorescent microscope and image algorithms are used to quantitate the amount of target proteins in different subcellular compartments. Measurements include endpoints such as the amount of a specific phosphorylated protein in the nucleus, translocation of transcription factors from cytoplasm to nucleus, the mitochondrial membrane potential, the shape of the nucleus, the length and number of neurites from neuronal cell cultures, the degree of microtubule dissociation, etc. Multiplexing is feasible (up to four separate channels) so that levels of several proteins and subcellular compartments can be measured simultaneously.

Genomics: Whole-Genome Methods

The three "omics" technologies (genomics, proteomics, and metabolomics) have all been used in toxicology research, although whole-genome microarray work is the one most commonly used and will be the focus here. In this approach, total mRNA is extracted from cells, converted into complementary DNA (cDNA), and quantified gene by gene. This is done either by hybridizing the cDNA to whole-genome chips, where the spots with cDNA are then fluorescently labeled, or by using sequencing techniques (e.g., "next-generation sequencing"). The use of genomics in toxicology is typically called toxicogenomics (Blomme, Yang, and Waring, 2009; Fielden and Zacharewski, 2001; Hamadeh et al., 2002a, b; Nuwaysir et al., 1999; Zhou et al., 2009), and in this case, one usually compares profiles of genes between treated and untreated samples. The absolute value of expression for any given gene is often not of interest, but instead, one focuses on the change in expression that is driven by the treatment.

These methods are technically challenging, partly because the data is somewhat noisy (both from a biological and a technical standpoint). Often, it is necessary to run multiple technical and biological replicates and to do appropriate averaging. A second challenge arises from the need to analyze data on 20 K or more genes. The first step of analysis is to do background subtraction and normalization. Then one takes differences between treatment and control samples and looks for genes that are differentially expressed in a statistically significant way. Multiple analysis strategies have been published to derive the differentially expressed genes (DEGs) and to then make sense out of the patterns. A common approach is to then map sets of DEGs to pathways or Gene Ontology (GO) processes, which have ties back to the biological literature. A consortium effort called the Microarray Quality Control (MAQC) consortium has held workshops and published analyses on reproducibility of results across platforms, labs, and time (MAQC-I) (Shi et al., 2006) and compared different strategies for analyzing case–control genomics data sets (MAQC-II) (Shi et al., 2010).

The advantage of this approach is that it provides a global, hypothesis-free view of the effect of a chemical. The disadvantage is that these analyses are expensive to run, making it difficult to perform full concentration–response analysis over many chemicals, time points, and cell systems. Also, the volume of data is so large that there are significant interpretation challenges. Another fundamental limitation of all RNA analysis techniques is that RNA levels (or their changes) are not well correlated with the corresponding protein levels (Gygi et al., 1999). For these reasons, microarray data is more often used for targeted studies to test a specific hypothesis on a limited number of chemicals.

In Vitro Pharmacokinetics

The preceding methods are used to understand the pharmacodynamic (PD) effects of a chemical, namely, what effect does it have on the cell it targets? Equally important are pharmacokinetics (PK), which determines how much of a parent molecule and/or its metabolites gets to the cell in the first place. Classical pharmacokinetic experiments are carried out in whole animals, in parallel with traditional whole-animal toxicity tests. But in the same way that *in vitro* tests are allowing rapid and inexpensive initial evaluation of chemical effects, corresponding high-throughput *in vitro* PK assays are becoming available for use.

Here, we describe two approaches that are being developed and are beginning to be applied in toxicology research. In the first case, the goal is to predict the oral dose of a chemical that would produce an effect at the level of a cell, a method sometimes called Reverse Toxicokinetics (RTK) or reverse dosimetry (Jamei et al., 2009a, b). *In vitro* HTS data can tell us what concentration at the cell is required to activate a pathway or inhibit an enzyme. The goal is then to predict what dose a human or animal would have to consume in order to reach an internal concentration equal to this activating value. If one makes some simplifying assumptions (including steady-state exposure, 100% oral bioavailability, 100% renal excretion, and that target-site concentration equals plasma concentration), there are two main experimental values one needs to measure. These are the intrinsic clearance rate (the rate at which the liver metabolizes the parent compound, which can be measured in primary hepatocytes for the species of interest, including humans) and plasma protein binding (which can again be measured in species-specific plasma). The rate of renal excretion is approximated as fraction unbound times the normal adult glomerular filtration rate (Rule et al., 2004). For each chemical of interest, one needs an analytical chemistry protocol to measure concentration of the parent in the medium being studied using liquid chromatography/mass spectrometry (LC/MS) or gas chromatography/mass spectrometry (GC/MS). These measurements also need to be run in two or more concentrations to make sure that the system is not

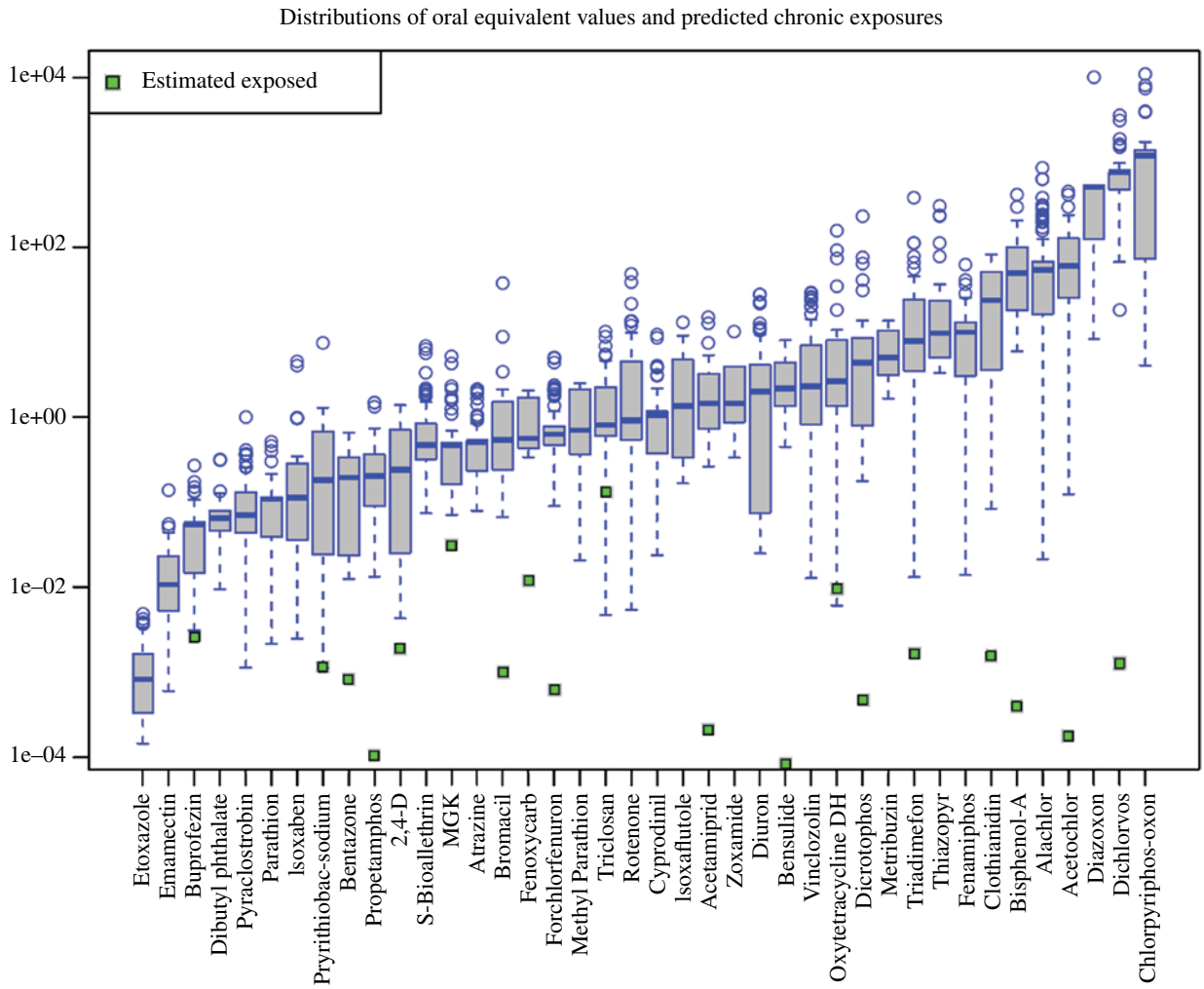


FIGURE 20.3 Results of RTK analysis for 40 chemicals. For each chemical, the plot shows the range of oral equivalent values as a box and whisker plot, with values of mg/kg/day. The expected chronic exposure for each chemical, due to pesticide residues in food, is shown by a green square. These chronic exposure values are derived for the most sensitive subpopulation (see Rotroff et al. for details). *Source:* Reproduced with permission from Rotroff et al. (2010). © 2010 Oxford University Press.

saturated. Finally, these experimental values can be used to parameterize a PK model, which in turn predicts a dose-to-concentration scaling factor, which is the concentration at steady state/dose rate (C_{ss}/DR). One then arrives at a prediction of the dose required to activate a pathway by dividing the IC50 (or similar quantity, in μM) by C_{ss}/DR. More complicated physiologically based pharmacokinetics models (PBPK) (Clewell, 1995; Krewski et al., 1994) could be used, but these typically require many more chemical-specific parameters, which are difficult to generate in a high-throughput manner. Without resorting to a full PBPK model, one can get some tissue-specific activating dose values by using QSAR models to predict certain partition coefficients, for instance, that for the blood–brain barrier (Schrodinger, 2010). The assumption of 100% oral bioavailability can also be relaxed and replaced with predictions for oral absorption from QSAR models (Schrodinger, 2010). Similar models can be built for other types of exposure (e.g., dermal or inhalation). Another extension that is possible is to use the PK model to predict peak exposure levels, for instance, to model the response to acute doses.

An example of this approach has been published by Rotroff et al. (2010). These authors carried out RTK analysis on a set of 40 chemicals, most of which were pesticide active ingredients. For all of these, a large collection of *in vitro* assay measurements were available, quantified as AC50s (concentration at which 50% of maximal activity was seen). For each chemical, the AC50 values were transformed into corresponding steady-state oral doses ("oral equivalent values") required to activate the assay. The results of this experiment are summarized in Figure 20.3, which shows the distribution of oral equivalent values for each of the *in vitro* assays. The most notable point is that the combination of PD and PK spread the dose range of effects over many orders of

magnitude from one chemical to another. At the extremes, etoxazole is expected to have *in vivo* biological activity at doses as low as 0.001 mg/kg/day, while no activity is expected for dichlorvos at doses below about 100 mg/kg/day. Recall that most of the assay data and all of the RTK parameters are based on human cells. One final point from Figure 20.3 is that the results of the RTK analysis can be directly compared with expected chronic exposures, in this case mainly due to pesticide residues in food. In the figure, the chronic exposure values are indicated by the green squares. From this figure, one sees that for most of these chemicals, there is a significant safety margin between the lowest oral equivalent value and the expected exposure. Unfortunately, for chemicals having no green squares, no systematic data on estimated exposure is currently available, so no such evaluation of the safety margin can be made.

The second aspect of PK that CT attempts to deal with is the understanding of what metabolites are produced from a parent molecule and what their relative effects are. The prediction of metabolites is principally done today using software systems such as Meteor, which uses a set of rules governing the types of reactions that can be catalyzed by xenobiotic-metabolizing enzymes. Although this process can be helpful in a qualitative way in identifying potentially dangerous metabolites, no information is available on relative rates of production of metabolites along various reaction branches.

A separate set of experimental approaches attempt to produce metabolites *in situ* within an assay system and so allow the measurement of the activity of the parent and all of its metabolites. A variety of ways have been proposed to do this, including running assays in primary hepatocytes or metabolically competent cell lines such as HepaRG (Aninat et al., 2006; Guillouzo et al., 2007). Another approach is to preincubate the chemical in a metabolically competent solution and then transfer the reaction mix to an assay system. This is a well-studied approach used in genotoxicity assays, using the S9 fraction, but this has proven difficult to scale up to high-throughput assays. A high-throughput analog of this approach has been developed in which the preincubation media is a solution containing immobilized enzymes (Lee et al., 2008).

20.2 DATA, DATABASES, KNOWLEDGE BASES, AND ONTOLOGIES

One aspect of CT is a focus on analyzing trends across many chemicals and many levels of biological organization, by using large data sets from many sources. This is the realm of bioinformatics, which integrates and analyzes large biological data sets. This section discusses some of the key data resources used in CT as well as approaches for developing custom databases. Two aspects of data are highlighted. The first is data itself and the second involves controlled vocabularies. This second piece is especially important when one is compiling large data sets to use for global analysis. If one source uses one term for a disease and another source uses another term, one cannot readily combine the two. The same is true for combining information on chemicals or genes, all of which have multiple names or identifiers that have to be reconciled before combining.

Chemicals

There are several large compilations of information on chemicals with structures, properties, and links to other data that is useful for CT research. These include DSSTox (Richard et al., 2006; U.S. EPA, 2008), which curates chemical structures and data sets useful for QSAR modeling; ACToR (Judson et al., 2008; U.S. EPA, 2008), which contains a large compilation of publicly accessible structure, property, *in vitro*, and toxicity data on environmental chemicals; ChemSpider (2008), which curates chemical structures and links to properties and a limited set of other data; and PubChem (NCBI, 2008), which links chemical structures to HTS *in vitro* data. Another useful chemical database is the Comparative Toxicogenomics Database (CTD) (Davis et al., 2009; Mattingly et al., 2006), which contains chemical–gene and chemical–disease links and will be discussed more in the following.

Genes

Genes provide the principal link between assays, pathways, and higher levels of effects of chemicals. Information on genes for humans and other species is compiled in several central databases including Entrez Gene (http://www.ncbi.nlm.nih.gov/gene) and GeneCards (http://www.ncbi.nlm.nih.gov/gene). Both of these provide definitive gene symbols and other IDs and links to the corresponding proteins, pathways, and gene-related diseases. Entrez Gene (as well as all of the NCBI Entrez databases) can be downloaded in their entirety and used to build local databases and applications.

Pathways

While we tend to measure the interaction of a chemical with a particular protein (for instance, in a binding assay) or the differential regulation of one or a few genes (for instance, in a transactivation assay), the toxicological consequence is mediated through a more complex pathway. The exact nature of this pathway modulation is rarely well understood, but this is one of the main goals of CT. Therefore, one will often organize the data from a series of assays into a picture of how a chemical is perturbing a pathway. With whole-genome microarray experiments, this is a standard analysis process, while with HTS assays, this approach is just becoming common.

Several commercial vendors of pathway data exist, but there are also widely used public repositories, including Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa, 2002; Kanehisa et al., 2002) and Pathway Commons (2010). Pathguide (http://www.pathguide.org/) provides a more complete listing of pathway resources.

Disease Links to Chemicals and Genes

Information on disease or toxicity can be broken into two areas. The first is actual links between chemicals and disease or between genes and disease. The second covers the terminology or vocabulary used to describe diseases, which will be discussed in the section on taxonomies. The largest compilation of data on chemical–gene and chemical–disease links curated from the literature is the CTD (Davis et al., 2009; Mattingly et al., 2006). Much of the chemical–gene data is derived from microarray experiments in which a chemical was shown to up- or downregulate specified genes. CTD also contains chemical–protein binding information. Another large source of chemical–gene data (indirectly) is PubChem, which holds data from a large number of HTS projects where chemicals have been screened through *in vitro* assays (NCBI, 2008). The Online Mendelian Inheritance in Man (OMIM) database (NCBI, 2009) is a compilation of information linking genes to disease. The origin of this was single gene diseases, but the coverage has subsequently broadened. Other sources include GeneCards, the Human Gene Mutation Database (HMGD) (http://www.hgmd.cf.ac.uk/ac/index.php), GenAtlas (http://www.genatlas.org/), and the Genetic Association Database (http://geneticassociationdb.nih.gov/).

In Vivo Toxicology Data

As we will discuss later, one important CT application is the development of models that link activity seen *in vitro* (genes, pathways) with whole-animal toxicity. We have shown how to measure activity against pathways for a set of chemicals using *in vitro* assays. During the model-building stage, one needs a set of chemicals for which both *in vitro* and high-quality *in vivo* toxicity data are available. Once a model is built and tested on this training set, it can be used to make predictions on chemicals for which only *in vitro* data is available. Ideally, the training set *in vivo* data will come from a uniform series of studies so that the results across the chemicals are comparable in all aspects. Additionally, it is best to have endpoints measured that are detailed enough to be mechanistically linked to the pathways one is measuring *in vitro*. For instance, an LD50 measurement is not particularly useful because many pathways can lead to death of the animal. An endpoint as specific as *in vivo* cholinesterase inhibition is ideal because it is linked directly to interactions of a chemical with the acetylcholinesterase protein. There are only a few compilations of *in vivo* data across many chemicals using uniform study protocols, the largest of which are ToxRefDB (Knudsen et al., 2009; Martin et al., 2009; U.S. EPA, 2008), developed by the EPA and mainly containing data on pesticides, and the National Toxicology Program's (NTP) database, which has information on a broader set of chemicals (NTP, 2011). Both of these resources provide detailed endpoint data at the dose-group level. Other high-quality data sets include the EPA Integrated Risk Information System (IRIS) (U.S. EPA, 2008), Agency for Toxic Substances and Disease Registry (ATSDR, 2011), and International Agency for Research on Cancer (IARC) (2011). These are largely compiled from the open literature for chemicals not covered by ToxRefDB or NTP.

Identifiers and Ontologies

A critical issue with building large compilations of data suitable for modeling is the need to have well-defined and consistent terminology. A pair of examples will illustrate this issue. First, as discussed earlier, chemicals can be identified through a variety of identifiers, including CASRN, EC number, common name, systematic name, or structure (in various formats). When pulling data on chemicals from reports or papers, some combination of these will be used and will often have small errors that have to be corrected. Hand-curated data efforts such as DSSTox will find and correct many errors and map all chemicals to unique IDs that allow comparison between chemicals and data sets. A second example arises from looking at pathology data. Even when compiling a cross-chemical data set from guideline studies, the terminology for describing specific lesions can differ from lab to lab and year to year. This drives the need for developing controlled vocabularies or taxonomies onto which all of the source data can be mapped.

We have already discussed the mapping of chemical data by identifiers such as CASRN and SMILES. Disease terminology is more complex because it involves both synonyms and hierarchies. For instance, hemangiosarcoma and angiosarcoma are synonyms, while both are sarcomas. A synonym for sarcoma is connective and soft tissue neoplasm, which is a cancer. Figure 20.4 shows a piece of the human disease ontology, found at http://diseaseontology.svn.sourceforge.net/viewvc/diseaseontology/trunk/HumanDO.obo, which is one of many ontologies that are contained in the Open Biological and Biomedical Ontologies (OBO) (Smith et al., 2007). By mapping all disease or pathology data from toxicology studies to a common ontology, one can more easily look for trends across chemicals and studies but can also conduct analyses across different levels of detail. A major reason for this has to do with statistical power. For instance, one might see one or two chemicals with each of several different diseases low in the hierarchy, which would not

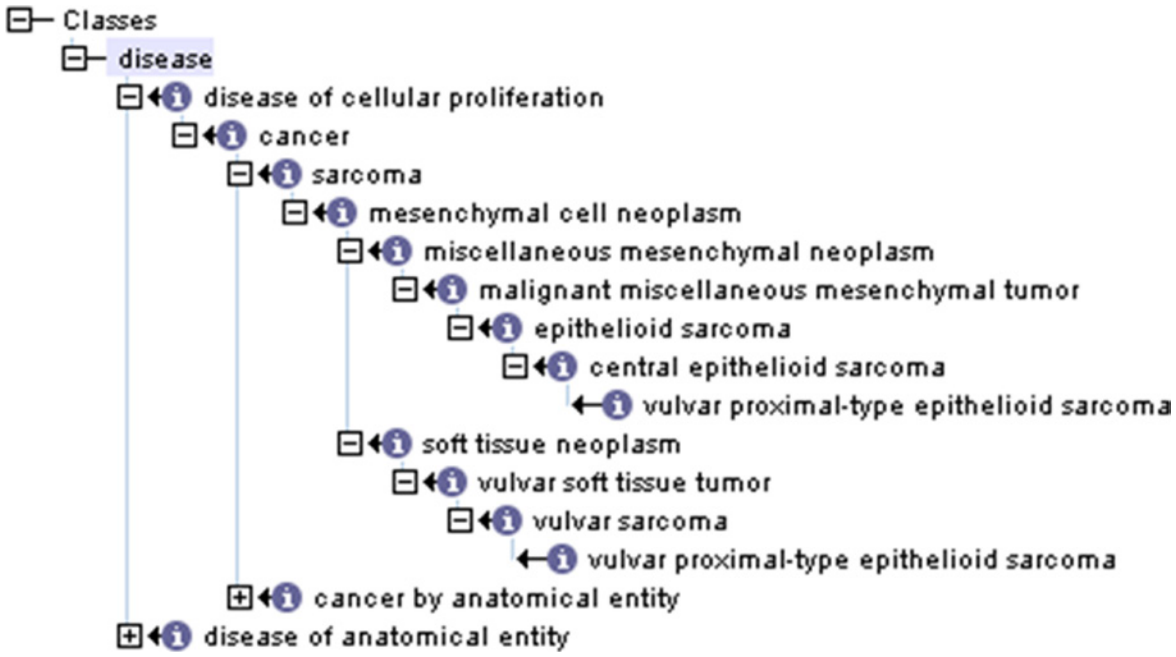


FIGURE 20.4 Subset of the human disease ontology.

provide enough statistical power to detect trends on a specific disease level. However, by combining these at a higher level in the hierarchy, trends can be observed. An ontology specifically related to chemical toxicology is ToxML (Richard et al., 2008), which is developing controlled vocabularies in areas such as genetic toxicity, carcinogenicity, developmental toxicity, neurotoxicity, and chronic toxicity.

Databases and Knowledge Bases

In order to perform large-scale analyses comparing data across chemicals, diseases, and pathways, it will often be most efficient to develop custom databases to organize all of the data prior to analyses. While this is not intended to be an introduction to database technology, a few definitions are worth discussing. First, a database generally refers to a "relational database," which is organized around a set of concepts and their relationships. Concepts might be chemicals, diseases, or genes. A chemical may have many properties and each property may have one or many values. For instance, properties of a chemical might be a molecular weight and a name. There will be a single molecular weight, but many names. Data is held in the database in tables that are functionally identical to spreadsheet tables, such as Excel (www.microsoft.com), but are linked with a set of keys. Figure 20.5 shows an example of two database tables that could manage parts of the chemical description information. Relational databases can be commercial products, such as Microsoft Access (www.microsoft.com) or

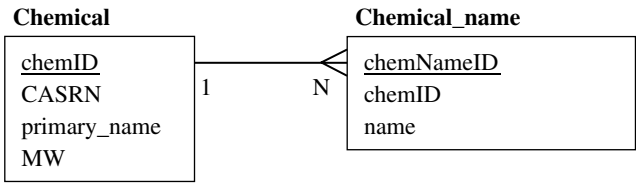


FIGURE 20.5 Example of two tables that could be part of a CT relational database, showing the variables or columns in the tables. Each type of object is defined by a unique ID and a set of properties (or columns). A chemical has a CASRN, a primary name, and a molecular weight, while a chemical name has its own ID and the string that holds the name itself. The crow's foot symbol connecting the two tables indicates that for each chemical ID (which defines a unique chemical), there can be many names (a one-to-many relationship). Each table contains a "primary key" (chemID and chemNameID), which are unique in a given table. Tables are linked through adding the primary key of one table to another table, in which case it takes on the role of a "foreign key," pointing back to the parent table.

Oracle (www.oracle.com), or open source applications, such as MySQL (www.mysql.com). In all cases, one manipulates the data in the database, including import and export, using languages generically called SQL for Structured Query Language.

A knowledge base extends the ideas behind relational databases to include more complex types of data and relationships between data. In particular, a knowledge base uses ontologies to define types of relationships, including

_chemical_arsenic	is_a	'chemical'
_chemical_arsenic	has_name	'arsenic'
_chemical_arsenic	causes	_tumor_101
_tumor_101	is_in_organ	'bladder'
_tumor_101	is_in_species	'human'
_tumor_101	has_exposure_route	'drinking water'

FIGURE 20.6 Example of the triples that could represent the statement "arsenic in drinking water causes bladder tumors in humans."

hierarchies, synonyms, and others. A simple relational implementation of a knowledge base can be created using a single database table called a triple store, which contains a set of subject–predicate–object triples (e.g., "arsenic causes bladder tumors"). In reality, such relationships contain modifying information, so a more complete statement might be: "Arsenic in drinking water causes bladder tumors in humans." "Arsenic" and "tumors" are the subject and object, and "causes" would be the predicate. In order to capture the more complex relationship in triples, one could create the following set of triples (Figure 20.6).

In this example, we use a set of variables that are represented by names starting with an underscore (e.g., _tumor_101). The reason for using these variables is to allow for complex modifiers to relationships, such as specifying the organ, the species, and the exposure route in the example earlier. Segaran et al. provide an easy-to-follow introduction to knowledge bases, triple stores, and other concepts mentioned here in the context of the semantic web (Segaran et al., 2009). The real power of this approach comes from being able to codify information coming from many different sources, often as free text in the open literature, in a way that computer models can be built. One important application is to import all of the open literature (e.g., from PubMed) into a knowledge base by parsing all sentences containing a set of predefined objects or terms, which might be chemicals, genes, and diseases as defined by common ontologies. An ontology of standardized relationships (including synonyms and hierarchy) would also have to be used. From this, we would find facts of the sort "chemical A causes disease B" or "gene X is downregulated in disease Y," with particular modifiers. In the original literature, these facts could have been stated many different ways, using a variety of synonyms. Once they are parsed into the knowledge base and standardized, the similarities are clear. In this way, one can see how many experiments support a particular statement, and in what circumstances. An extension of this idea comes from finding statements of the sort "A causes B" and "B causes C," each with significant support. From this, one can make the inference that "A causes C," even in the absence of any direct evidence. Such inferences form hypotheses that can then be tested experimentally. Significant effort is going into the development of ontologies across the biological sciences, and these are enabling the construction of biologically and toxicologically oriented knowledge bases. For an introduction to this field, one should consult the home page of the OBO Foundry (http://www.obofoundry.org/).

20.3 EXAMPLE CT APPLICATIONS

In this section, we briefly describe several current or proposed applications of CT methods.

Predicting Hazard from *In Vitro* Data: Toxicity Signatures

In general, a signature is a set of biomarkers whose collective activity is predictive of some measureable outcome. There are many examples of predictive signatures in the fields of genetics and genomics, some of which are now being translated into clinically useful applications for predicting patient responses to pharmaceutical agents (Buyse et al., 2006; Klein et al., 2009). A toxicity signature, as used here, models the association between a set of *in vitro* predictors (i.e., assay results) and a given toxicological endpoint or hazard. Thus, the purpose of a toxicity signature is to enable prediction of a particular hazard from *in vitro* data used to develop that signature.

An illustrative example of using *in vitro* data to predict hazard is provided by Martin et al. (2011). In this study, the *in vitro* (predictor) data were results from over 500 biochemical and cell-based assays that had been run across 309 chemicals as part of the U.S. EPA's ToxCast program (Dix et al., 2007; Judson et al., 2010). The hazard data were results from multigenerational reproductive studies, as captured in the aforementioned ToxRefDB (see Section "*In Vivo* Toxicology Data"). A subset of the 309 chemicals having both *in vitro* ToxCast data and high-quality reproductive hazard data were used to build a predictive signature of rodent reproductive toxicity. This signature was empirically derived by mining the statistical associations between *in vitro* assay results and positive reproductive toxicants. It predicted reproductive hazard from a combination of assays probing nuclear receptors, cytochrome P450 activity, G protein-coupled receptors, and cell signaling pathways. There is great potential utility toward risk assessment in a signature built upon the results of rapid, high-throughput *in vitro* assays. The immediate utility of such a signature lies in the mechanistic information it provides for additional, targeted testing strategies.

Tiered Testing and Prioritization

One of the first applications of CT approaches in the regulatory arena will be screening for chemicals that have the potential to cause toxicity through modes of action identified using *in vitro* assays. This is similar to current QSAR or expert systems approaches that create structure alerts, that is, a chemical contains a structural feature that is common among chemicals that are mutagens or cause other types of toxicity. Two widely used structure-based methods are Derek and Oncologic (Dearden, 2003; Patlewicz et al., 2003). An example being developed at the EPA is the use of assays for estrogen (ER), androgen (AR), and other endocrine-relevant receptor interactions that can be used to prioritize chemicals going into the EPA Endocrine Disruption Screening Program (EDSP). There are over 4000 compounds (primarily pesticide active and inert ingredients) that are required to go through tier 1 of this program, at an estimated cost of approximately $0.5M per chemical (U.S. EPA, 2007). In addition to the high cost, another issue with the EDSP is that the throughput is limited to about 100 chemicals per year by a lack of testing labs and evaluation resources at EPA. Therefore, it would be advantageous to use quick, inexpensive HTS assays to evaluate many of the candidate chemicals and prioritize the ones with ER, AR, or other endocrine-relevant activity to be run first. This situation in which there are many chemicals to be evaluated, yet few have complete data, is common in the environmental arena, including among industrial chemicals and chemicals found in surface and/or drinking water.

Formal decision analysis approaches are being developed to support chemical prioritization needs in the United States and Europe (Gabbert and Weikard, 2010). The Toxicological Prioritization Index (ToxPi) framework is one such approach that can be tailored to diverse sets of chemicals and prioritization tasks (Reif et al., 2010). ToxPi provides a visual, weight-of-evidence index that can be used to rank and compare chemicals. Its initial application was to aid in the prioritization of chemicals for putative endocrine activity, in support of the EDSP. As in Figure 20.7, composite activity scores (profiles) across sets of endocrine-relevant data (slices) for each chemical were calculated to rank all 309 ToxCast phase I chemicals from the highest (upper right) to the lowest (lower left) priority. In the face of practical temporal and economic limitations, this estimate of potential endocrine activity

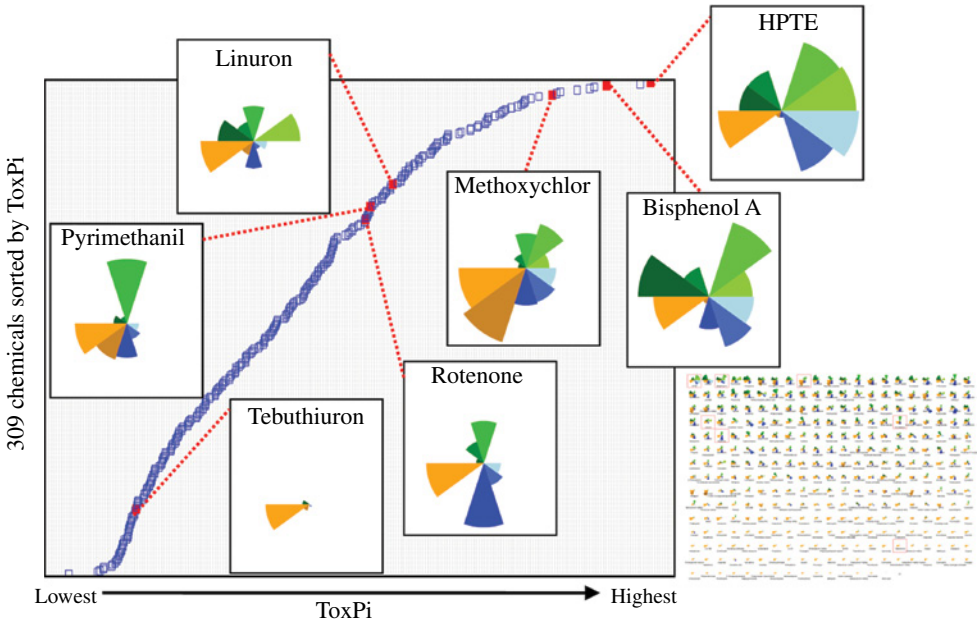


FIGURE 20.7 ToxPi prioritization of all 309 ToxCast phase I chemicals. The ToxPi (horizontal axis) for each chemical (vertical axis) is symbolized by a blue box, sorted according to overall ToxPi. Profiles for seven reference chemicals are shown, where green slices indicate *in vitro* assays, orange slices indicate selected chemical properties, and blue slices indicate biologically relevant pathway scores. From the highest ToxPi to the lowest, the reference chemicals are HPTE (a metabolite of methoxychlor), BPA, methoxychlor, linuron, pyrimethanil, rotenone, and tebuthiuron. The inset shows a grid of the sorted ToxPi profiles for all chemicals, with reference chemicals highlighted by solid red boxes. *Source:* Reproduced with permission from Reif et al. (2010).

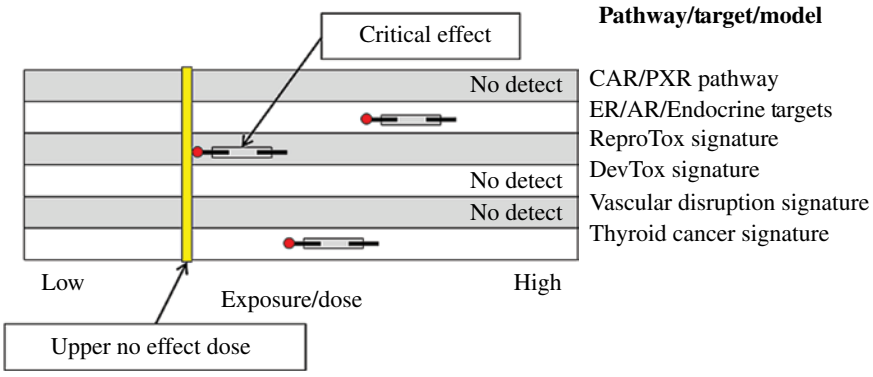


FIGURE 20.8 Illustration of a chemical report card for a chemical tested against several pathways.

provides a formal rationale for prioritizing resources toward further testing. Alternatively, the ToxPi profiles could be used for chemical "read-across," analogous to QSAR structural alert models. The read-across can be implemented in terms of overall ToxPi bins/clusters of chemicals having similar profiles, or subsets of slices can be interpreted as *in vitro* "alerts" to support targeted testing decisions.

High-Throughput Risk Assessment

Our final application is one we call "High-Throughput Risk Assessment" (HTRA) (Judson et al., 2011). HTRA provides first-order estimates of maximum doses of a chemical that cause no biological activity against specified biological pathways. The assumption is that if a pathway is not perturbed, then no adverse effects can be caused. HTRA uses a combination of HTS assay data, which provides the concentration (as the AC50 in μM) at which a chemical perturbs a target or pathway, plus *in vitro* estimates of PK using the RTK method that was described earlier. One can then derive an estimate of the oral steady-state dose at which a chemical will perturb the pathway, given by

$$\text{Biological pathway altering dose}\left(\text{BPAD}\right)=\frac{\text{AC50}}{\left(\text{C}_{ss}/\text{DR}\right)}$$

HTRA then subjects these quantities to estimates of population variability and uncertainty to yield confidence intervals about the perturbing dose. A lower end of the confidence interval can then be taken as an estimate of an upper safe dose of the chemical, in the absence of other available information. For each chemical, one will need to test against each of multiple toxicity pathways or signatures. Figure 20.8 conceptually illustrates the results of such an analysis in the form of an HTRA Report Card. Each pathway will be tested using one or more assays, which produce an AC50 value. These are transformed into BPAD (dose) values using the aforementioned equation, and confidence intervals are built. The low-dose end of the confidence intervals (annotated with the red dot) is the lowest BPAD or BPADL. The pathway producing the lowest BPADL is then the *in vitro* critical effect, in analogy to the critical effects derived from a traditional animal-based toxicity study. This value can then be compared with actual exposure values, and to first order, we can assume that exposure below this value will not be harmful. This approach is very new and many open issues remain, including the best ways to estimate uncertainty and population variability and the need to link *in vitro* pathway perturbations to *in vivo* adverse effects.

20.4 SUMMARY

CT is an emerging field that (as we define it here) combines HTS data on chemicals with information from the literature on genes, pathways, diseases, and toxic effects. These combined data sets are used for a variety of purposes ranging from understanding the mechanistic basis of chemical toxicity to screening and prioritization of chemicals for further testing. The world of CT is summarized in Figure 20.9, where the center section shows the types of information, data, and tests that are brought to bear. In vitro assays play a central role, and the wide variety of assays used is exemplified at the right of the figure. A second key piece of CT is the integration of many types of data, in flexible ways, which derives the need to develop databases, knowledge bases, and ontologies, shown at the left of the diagram. Finally, as shown at the bottom, there are an emerging set of tools and applications.

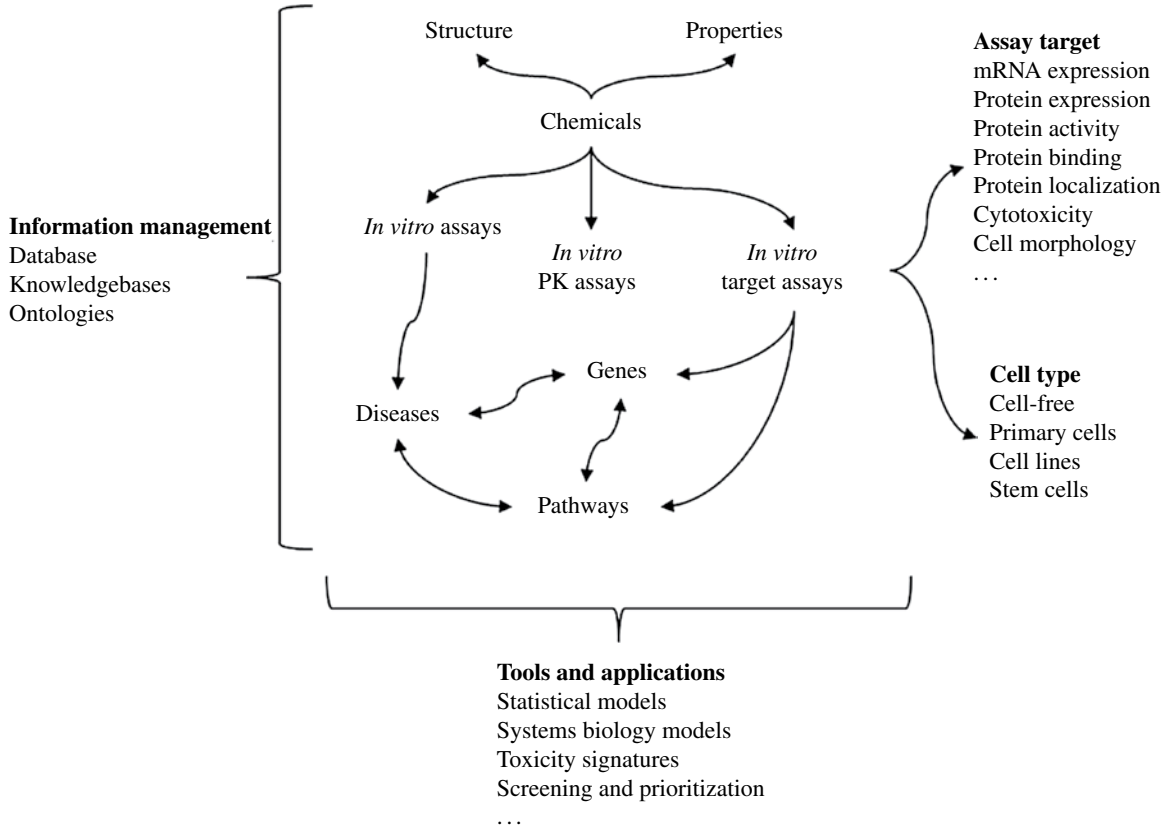


FIGURE 20.9 Summary of linkages between types of computational toxicology data and applications described in this chapter.

SUGGESTED READING

Aninat C, Piton A, Glaise D, Le Charpentier T, Langouët S, Morel F, Guguen-Guillouzo C, Guillouzo A. Expression of cytochromes P450, conjugating enzymes and nuclear receptors in human hepatoma HepaRG cells. Drug Metab Dispos 2006;34(1):75–83.

ATSDR. 2011. Agency for toxic substances and disease registry, toxic substances portal. Available at http://www.atsdr.cdc.gov/toxfaqs/index.asp. Accessed April 22, 2011.

Berg EL, Yang J, Melrose J, Nguyen D, Privat S, Rosler E, Kunkel EJ, Ekins S. Chemical target and pathway toxicity mechanisms defined in primary human cell systems. J Pharmacol Toxicol Methods 2010;61(1):3–15.

Bleicher KH, Bohm HJ, Müller K, Alanine AI. Hit and lead generation: beyond high-throughput screening. Nat Rev Drug Discov 2003;2(5):369–378.

Blomme EA, Yang Y, Waring JF. Use of toxicogenomics to understand mechanisms of drug-induced hepatotoxicity during drug discovery and development. Toxicol Lett 2009;186(1):22–31.

Boess F, Kamber M, Romer S, Gasser R, Muller D, Albertini S, Suter L. Gene expression in two hepatic cell lines, cultured primary hepatocytes, and liver slices compared to the in vivo liver gene expression in rats: possible implications for toxicogenomics use of in vitro systems. Toxicol Sci 2003;73(2):386–402.

Bullen A. Microscopic imaging techniques for drug discovery. Nat Rev Drug Discov 2008;7(1):54–67.

Buyse M, Loi S, van't Veer L, Viale G, Delorenzi M, Glas AM, d'Assignies MS, Bergh J, Lidereau R, Ellis P, Harris A, Bogaerts J, Therasse P, Floore A, Amakrane M, Piette F, Rutgers E, Sotiriou C, Cardoso F, Piccart MJ, TRANSBIG Consortium. Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. J Natl Cancer Inst 2006;98(17):1183–1192.

ChemSpider. 2008. ChemSpider: Building a structure centric community for chemists. Available at http://www.chemspider.com/. Accessed August 8, 2008.

Clewell HJ 3rd. The application of physiologically based pharmacokinetic modeling in human health risk assessment of hazardous substances. Toxicol Lett 1995;79(1–3):207–217.

Davis AP, Murphy CG, Saraceni-Richards CA, Rosenstein MC, Wiegers TC, Mattingly CJ. Comparative Toxicogenomics Database: a knowledgebase and discovery tool for chemical-gene-disease networks. Nucleic Acids Res 2009;37(Database issue):D786–D792.

Daylight. 2008. SMILES (Simplified Molecular Input Line Entry System). Available at http://www.daylight.com/dayhtml/doc/theory/theory.smiles.html. Accessed October 24, 2008.

Dearden JC. In silico prediction of drug toxicity. J Comput Aided Mol Des 2003;17(2–4):119–127.

Dick E, Rajamohan D, Ronksley J, Denning C. Evaluating the utility of cardiomyocytes from human pluripotent stem cells for drug screening. Biochem Soc Trans 2010;38(4):1037–1045.

Dix DJ, Houck KA, Martin MT, Richard AM, Setzer RW, Kavlock RJ. The ToxCast program for prioritizing toxicity testing of environmental chemicals. Toxicol Sci 2007;95(1):5–12.

Fielden MR, Zacharewski TR. Challenges and limitations of gene expression profiling in mechanistic and predictive toxicology. Toxicol Sci 2001;60(1):6–10.

Gabbert S, Weikard H-P. A theory of chemicals testing and regulation. Nat Resour Forum 2010;34(2):155–164.

Giuliano KA, Cheung WS, Curran DP, Day BW, Kassick AJ, Lazo JS, Nelson SG, Shin Y, Taylor DL. Systems cell biology knowledge created from high content screening. Assay Drug Dev Technol 2005;3(5):501–514.

Giuliano KA, Johnston PA, Gough A, Taylor DL. Systems cell biology based on high-content screening. Methods Enzymol 2006;414:601–619.

Guillouzo A, Corlu A, Aninat C, Glaise D, Morel F, Guguen-Guillouzo C. The human hepatoma HepaRG cells: a highly differentiated model for studies of liver metabolism and toxicity of xenobiotics. Chem Biol Interact 2007;168(1):66–73.

Gygi SP, Rochon Y, Franza BR, Aebersold R. Correlation between protein and mRNA abundance in yeast. Mol Cell Biol 1999; 19(3):1720–1730.

Hamadeh HK, Bushel PR, Jayadev S, DiSorbo O, Bennett L, Li L, Tennant R, Stoll R, Barrett JC, Paules RS, Blanchard K, Afshari CA. Prediction of compound signature using high density gene expression profiling. Toxicol Sci 2002a;67(2):232–240.

Hamadeh HK, Bushel PR, Jayadev S, Martin K, DiSorbo O, Sieber S, Bennett L, Tennant R, Stoll R, Barrett JC, Blanchard K, Paules RS, Afshari CA. Gene expression analysis reveals chemical-specific profiles. Toxicol Sci 2002b;67(2):219–231.

Huang P, He Z, Ji S, Sun H, Xiang D, Liu C, Hu Y, Wang X, Hui L. Induction of functional hepatocyte-like cells from mouse fibroblasts by defined factors. Nature 2011;475:386–389.

IARC. (2011). International Agency on Research on Cancer. Available at http://www.iarc.fr/. Accessed April 22, 2011.

Inglese J, Auld DS, Jadhav A, Johnson RL, Simeonov A, Yasgar A, Zheng W, Austin CP. Quantitative high-throughput screening: a titration-based approach that efficiently identifies biological activities in large chemical libraries. Proc Natl Acad Sci U S A 2006;103(31): 11473–11478.

Inglese J, Johnson RL, Simeonov A, Xia M, Zheng W, Austin CP, Auld DS. High-throughput screening assays for the identification of chemical probes. Nat Chem Biol 2007;3(8):466–479.

IUPAC. 1993. IUPAC nomenclature of organic chemistry. Available at http://www.acdlabs.com/iupac/nomenclature/. Accessed April 20, 2011.

IUPAC. 2008. IUPAC—International Chemical Identifier. Available at http://old.iupac.org/projects/2000/2000-025-1-800.html. Accessed October 24, 2008.

Jamei M, Marciniak S, Feng K, Barnett A, Tucker G, Rostami-Hodjegan A. The Simcyp population-based ADME simulator. Expert Opin Drug Metab Toxicol 2009a;5(2):211–223.

Jamei M, Turner D, Yang J, Neuhoff S, Polak S, Rostami-Hodjegan A, Tucker G. Population-based mechanistic prediction of oral drug absorption. AAPS J 2009b;11(2):225–237.

Judson R, Richard A, Dix D, Houck K, Elloumi F, Martin M, Cathey T, Transue TR, Spencer R, Wolf M. ACToR—Aggregated Computational Toxicology Resource. Toxicol Appl Pharmacol 2008;233(1):7–13.

Judson RS, Houck KA, Kavlock RJ, Knudsen TB, Martin MT, Mortensen HM, Reif DM, Rotroff DM, Shah I, Richard AM, Dix DJ. Predictive *in vitro* screening of environmental chemicals—the ToxCast project. Environ Health Perspect 2010;118(4):485–492.

Judson RS, Kavlock RJ, Setzer RW, Hubal EA, Martin MT, Knudsen TB, Houck KA, Thomas RS, Wetmore BA, Dix DJ. Estimating toxicity-related biological pathway altering doses for high-throughput chemical risk assessment. Chem Res Toxicol 2011;24(4):451–462.

Kanehisa M. The KEGG database. Novartis Found Symp 2002;247:91–101; discussion 101–103, 119–128, 244–152.

Kanehisa M, Goto S, Kawashima S, Nakaya A. The KEGG databases at GenomeNet. Nucleic Acids Res 2002;30(1):42–46.

Klein TE, Altman RB, Eriksson N, Gage BF, Kimmel SE, Lee MT, Limdi NA, Page D, Roden DM, Wagner MJ, Caldwell MD, Johnson JA, International Warfarin Pharmacogenetics Consortium. Estimation of the warfarin dose with clinical and pharmacogenetic data. N Engl J Med 2009;360(8):753–764.

Knudsen TB, Martin MT, Kavlock RJ, Judson RS, Dix DJ, Singh AV. Profiling the activity of environmental chemicals in prenatal developmental toxicity studies using the U.S. EPA's ToxRefDB. Reprod Toxicol 2009;28:209–219.

Krewski D, Withey JR, Ku LF, Andersen ME. Applications of physiologic pharmacokinetic modeling in carcinogenic risk assessment. Environ Health Perspect 1994;102(Suppl 11):37–50.

Leadscope. 2010. Leadscope Model Applier. Available at http://www.leadscope.com/. Accessed August 7, 2014.

Lee MY, Kumar RA, Sukumaran SM, Hogg MG, Clark DS, Dordick JS. Three-dimensional cellular microarray for high-throughput toxicology assays. Proc Natl Acad Sci U S A 2008;105(1):59–63.

Lunn CA. Label-free screening assays: a strategy for finding better drug candidates. Future Med Chem 2010;2(11):1703–1716.

Martin MT, Knudsen TB, Reif DM, Houck KA, Judson RS, Kavlock RJ, Dix DJ. Predictive model of rat reproductive toxicity from ToxCast high throughput screening. Biol Reprod 2011;85:327–339.

Martin MT, Mendez E, Corum DG, Judson RS, Kavlock RJ, Rotroff DM, Dix DJ. Profiling the reproductive toxicity of chemicals from multigeneration studies in the toxicity reference database (ToxRefDB). Toxicol Sci 2009;110:181–190.

Mattingly CJ, Rosenstein MC, Davis AP, Colby GT, Forrest JN Jr, Boyer JL. The comparative toxicogenomics database: a cross-species resource for building chemical-gene interaction networks. Toxicol Sci 2006;92(2):587–595.

Mayr LM, Bojanic D. Novel trends in high-throughput screening. Curr Opin Pharmacol 2009;9(5):580–588.

NCBI. 2008. PubChem. Available at http://pubchem.ncbi.nlm.nih.gov/. Accessed August 8, 2008.

NCBI. 2009. *Online Mendelian Inheritance in Man, OMIM (TM). McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, MD) and National Center for Biotechnology Information, National Library of Medicine* (Bethesda, MD). Available at http://www.ncbi.nlm.nih.gov/omim/. Accessed October 26, 2009.

Noraberg J, Poulsen FR, Blaabjerg M, Kristensen BW, Bonde C, Montero M, Meyer M, Gramsbergen JB, Zimmer J. Organotypic hippocampal slice cultures for studies of brain damage, neuroprotection and neurorepair. Curr Drug Targets CNS Neurol Disord 2005;4(4):435–452.

NTP. 2011. National Toxicology Program. Available at http://ntp.niehs.nih.gov/. Accessed April 22, 2011.

Nuwaysir EF, Bittner M, Trent J, Barrett JC, Afshari CA. Microarrays and toxicology: the advent of toxicogenomics. Mol Carcinog 1999;24(3):153–159.

O'Shea SH, Schwarz J, Kosyk O, Ross PK, Ha MJ, Wright FA, Rusyn I. In vitro screening for population variability in chemical toxicity. Toxicol Sci 2011;119(2):398–407.

Parng C, Seng WL, Semino C, McGrath P. Zebrafish: a preclinical model for drug screening. Assay Drug Dev Technol 2002;1(1 Pt 1):41–48.

PathwayCommons. 2010. Pathway Commons. Available at http://www.pathwaycommons.org/pc/home.do. Accessed March 3, 2010.

Patlewicz G, Rodford R, Walker JD. Quantitative structure-activity relationships for predicting mutagenicity and carcinogenicity. Environ Toxicol Chem 2003;22(8):1885–1893.

Reif DM, Martin MT, Tan SW, Houck KA, Judson RS, Richard AM, Knudsen TB, Dix DJ, Kavlock RJ. Endocrine profiling and prioritization of environmental chemicals using ToxCast data. Environ Health Perspect 2010;118(12):1714–1720.

Richard AM, Gold LS, Nicklaus MC. Chemical structure indexing of toxicity data on the internet: moving toward a flat world. Curr Opin Drug Discov Devel 2006;9(3):314–325.

Richard AM, Williams CR. Distributed structure-searchable toxicity (DSSTox) public database network: a proposal. Mutat Res 2002;499(1):27–52.

Richard AM, Yang C, Judson RS. Toxicity data informatics: supporting a new paradigm for toxicity prediction. Toxicol Mech Methods 2008;18:103–118.

Rimsza LM, Leblanc ML, Unger JM, Miller TP, Grogan TM, Persky DO, Martel RR, Sabalos CM, Seligmann B, Braziel RM, Campo E, Rosenwald A, Connors JM, Sehn LH, Johnson N, Gascoyne RD. Gene expression predicts overall survival in paraffin-embedded tissues of diffuse large B-cell lymphoma treated with R-CHOP. Blood 2008;112(8):3425–3433.

Roberts RA, Sabalos CM, LeBlanc ML, Martel RR, Frutiger YM, Unger JM, Botros IW, Rounseville MP, Seligmann BE, Miller TP, Grogan TM, Rimsza LM. Quantitative nuclease protection assay in paraffin-embedded tissue replicates prognostic microarray gene expression in diffuse large-B-cell lymphoma. Lab Invest 2007;87(10):979–997.

Rotroff DM, Wetmore BA, Dix DJ, Ferguson SS, Clewell HJ, Houck KA, Lecluyse EL, Andersen ME, Judson RS, Smith CM, Sochaski MA, Kavlock RJ, Boellmann F, Martin MT, Reif DM, Wambaugh JF, Thomas RS. Incorporating human dosimetry and exposure into high-throughput in vitro toxicity screening. Toxicol Sci 2010;117(2):348–358. DOI: 10.1093/toxsci/kfq220.

Rule AD, Gussak HM, Pond GR, Bergstralh EJ, Stegall MD, Cosio FG, Larson TS. Measured and estimated GFR in healthy potential kidney donors. Am J Kidney Dis 2004;43(1):112–119.

Russom CL, Williams CR, Stewart TW, Swank AE, Richard AM. 2008, Updated version: EPAFHM_v4a_617_15Jun2007). DSSTox EPA Fathead Minnow Acute Toxicity Database (EPAFHM): SDF Files and Documentation. Available at http://www.epa.gov/ncct/dsstox/. Accessed August 8, 2008.

Schrodinger. 2010. QikProp. Available at https://www.schrodinger.com/products/14/17/. Accessed March 24, 2010.

Segaran T, Evans C, Taylor J. *Programming the Semantic Web*. Sebastopol (CA): O'Reilly; 2009.

Shi L, Campbell G, Jones WD, Campagne F, Wen Z, Walker SJ, Su Z, Chu TM, Goodsaid FM, Pusztai L, Shaughnessy JD Jr, Oberthuer A, Thomas RS, Paules RS, Fielden M, Barlogie B, Chen W, Du P, Fischer M, Furlanello C, Gallas BD, Ge X, Megherbi DB, Symmans WF, Wang MD, Zhang J, Bitter H, Brors B,Bushel PR, Bylesjo M, Chen M, Cheng J, Cheng J, Chou J, Davison TS, Delorenzi M, Deng Y, Devanarayan V, Dix DJ, Dopazo J, Dorff KC, Elloumi F, Fan J, Fan S, Fan X, Fang H, Gonzaludo N, Hess KR, Hong H, Huan J, Irizarry RA, Judson R, Juraeva D, Lababidi S, Lambert CG, Li L, Li Y, Li Z, Lin SM, Liu G, Lobenhofer EK, Luo J, Luo W, McCall MN, Nikolsky Y, Pennello GA, Perkins RG, Philip R, Popovici V, Price ND, Qian F, Scherer A, Shi T, Shi W, Sung J, Thierry-Mieg D,Thierry-Mieg J, Thodima V, Trygg J, Vishnuvajjala L, Wang SJ, Wu J, Wu Y, Xie Q, Yousef WA, Zhang L, Zhang X, Zhong S, Zhou Y, Zhu S, Arasappan D, Bao W,Lucas AB, Berthold F, Brennan RJ, Buness A, Catalano JG, Chang C, Chen R, Cheng Y, Cui J, Czika W, Demichelis F, Deng X, Dosymbekov D, Eils R, Feng Y,Fostel J, Fulmer-Smentek S, Fuscoe JC, Gatto L, Ge W, Goldstein DR, Guo L, Halbert DN, Han J, Harris SC, Hatzis C, Herman D, Huang J, Jensen RV, Jiang R, Johnson CD, Jurman G, Kahlert Y, Khuder SA, Kohl M, Li J, Li L, Li M, Li QZ, Li S, Li Z, Liu J, Liu Y, Liu Z, Meng L, Madera M, Martinez-Murillo F, Medina I,Meehan J, Miclaus K, Moffitt RA, Montaner D, Mukherjee P, Mulligan GJ, Neville P, Nikolskaya T, Ning B, Page GP, Parker J, Parry RM, Peng X, Peterson RL,Phan JH, Quanz B, Ren Y, Riccadonna S, Roter AH, Samuelson FW, Schumacher MM, Shambaugh JD, Shi Q, Shippy R, Si S, Smalter A, Sotiriou C, Soukup M,Staedtler F, Steiner G, Stokes TH, Sun Q, Tan PY, Tang R, Tezak Z, Thorn B, Tsyganova M, Turpaz Y, Vega SC, Visintainer R, von Frese J, Wang C, Wang E,Wang J, Wang W, Westermann F, Willey JC, Woods M, Wu S, Xiao N, Xu J, Xu L, Yang L, Zeng X, Zhang J, Zhang L, Zhang M, Zhao C, Puri RK, Scherf U, Tong W, Wolfinger RD; MAQC Consortium. The MicroArray Quality Control (MAQC)-II study of common practices for the development and validation of microarray-based predictive models. Nat Biotechnol 2010;28(8):827–838.

Shi L., Reid LH, Jones WD, Shippy R, Warrington JA, Baker SC, Collins PJ, de Longueville F, Kawasaki ES, Lee KY, Luo Y, Sun YA, Willey JC, Setterquist RA, Fischer GM, Tong W, Dragan YP, Dix DJ, Frueh FW, Goodsaid FM, Herman D, Jensen RV, Johnson CD, Lobenhofer EK, Puri RK, Schrf U,Thierry-Mieg J, Wang C, Wilson M, Wolber PK, Zhang L, Amur S, Bao W,

Barbacioru CC, Lucas AB, Bertholet V, Boysen C, Bromley B, Brown D, Brunner A,Canales R, Cao XM, Cebula TA, Chen JJ, Cheng J, Chu TM, Chudin E, Corson J, Corton JC, Croner LJ, Davies C, Davison TS, Delenstarr G, Deng X, Dorris D,Eklund AC, Fan XH, Fang H, Fulmer-Smentek S, Fuscoe JC, Gallagher K, Ge W, Guo L, Guo X, Hager J, Haje PK, Han J, Han T, Harbottle HC, Harris SC,Hatchwell E, Hauser CA, Hester S, Hong H, Hurban P, Jackson SA, Ji H, Knight CR, Kuo WP, LeClerc JE, Levy S, Li QZ, Liu C, Liu Y, Lombardi MJ, Ma Y,Magnuson SR, Maqsodi B, McDaniel T, Mei N, Myklebost O, Ning B, Novoradovskaya N, Orr MS, Osborn TW, Papallo A, Patterson TA, Perkins RG, Peters EH,Peterson R, Philips KL, Pine PS, Pusztai L, Qian F, Ren H, Rosen M, Rosenzweig BA, Samaha RR, Schena M, Schroth GP, Shchegrova S, Smith DD, Staedtler F, Su Z, Sun H, Szallasi Z, Tezak Z, Thierry-Mieg D, Thompson KL, Tikhonova I, Turpaz Y, Vallanat B, Van C, Walker SJ, Wang SJ, Wang Y, Wolfinger R, Wong A,Wu J, Xiao C, Xie Q, Xu J, Yang W, Zhang L, Zhong S, Zong Y, Slikker W Jr. The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. Nat Biotechnol 2006;24(9):1151–1161.

Smith B, Ashburner M, Rosse C, Bard J, Bug W, Ceusters W, Goldberg LJ, Eilbeck K, Ireland A, Mungall CJ; OBI Consortium, Leontis N, Rocca-Serra P, Ruttenberg A, Sansone SA, Scheuermann RH, Shah N, Whetzel PL, Lewis S. The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. Nat Biotechnol 2007;25(11):1251–1255.

Smith MV, Boyd WA, Kissling GE, Rice JR, Snyder DW, Portier CJ, Freedman JH. A discrete time model for the analysis of medium-throughput *C. elegans* growth data. PLoS One 2009;4(9):e7018.

Sundberg SA. High-throughput and ultra-high-throughput screening: solution- and cell-based approaches. Curr Opin Biotechnol 2000;11(1):47–53.

Taipa MA. Immunoassays: biological tools for high throughput screening and characterisation of combinatorial libraries. Comb Chem High Throughput Screen 2008;11(4):325–335.

The International HapMap Consortium. A haplotype map of the human genome. Nature 2005;437(7063):1299–1320.

U.S. EPA. 2007. Draft Information Collection Request (ICR): Tier 1 Screening of Certain Chemicals Under the Endocrine Disruptor Screening Program (EDSP). Available at http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPPT-2007-1081-0002. Accessed April 20, 2011.

U.S. EPA. 2008a. ACToR Online (Aggregated Computational Toxicology Resource). Available at http://actor.epa.gov/actor. Accessed November 19, 2008.

U.S. EPA. 2008b. EPA Integrated Risk Information System. Available at http://cfpub.epa.gov/ncea/iris/. Accessed October 24, 2008.

U.S. EPA. 2008c. TOXCST: Research Chemical Inventory for EPA's ToxCastTM Program: Structure-Index File. Available at http://www.epa.gov/ncct/dsstox/sdf_toxcst.html. Accessed August 8, 2008.

U.S. EPA. 2008d. ToxRefDB: EPA Toxicity Reference Database. Available at http://actor.epa.gov/toxrefdb/faces/Home.jsp. Accessed February 1, 2011.

U.S. EPA. 2010. EPI Suite v 4.5. Available at http://www.epa.gov/oppt/exposure/pubs/episuitedl.htm. Accessed March 24, 2010.

Westerink WM, Schoonen WG. Cytochrome P450 enzyme levels in HepG2 cells and cryopreserved primary human hepatocytes and their induction in HepG2 cells. Toxicol In Vitro 2007;21(8):1581–1591.

Zhou T, Chou J, Watkins PB, Kaufmann WK. Toxicogenomics: transcription profiling for toxicology assessment. EXS 2009;99:325–366.

21

EPIDEMIOLOGICAL ISSUES IN OCCUPATIONAL AND ENVIRONMENTAL HEALTH

MICHAEL S. BLOOM AND JOHN E. VENA

Environmental epidemiology has become perhaps one of the most challenging specialty areas in human health research, with a complex theoretical paradigm and causal pathways involved that necessitate a multidisciplinary approach with toxicologists partnering with epidemiologists. Therefore, toxicologists and epidemiologists must learn each other's discipline, become familiar with the special use of terms and concepts, understand these concepts, and have insights into the methodology used by each other's discipline. We must learn to appreciate and understand different approaches and scientific cultures and perspectives. In this chapter, we highlight the principles of the epidemiologic approach as it applies to environmental and occupational exposures and health, with attention to toxicologic applications in the field and with specific examples. More specifically, this chapter will address:

- Why toxicologists are critical to successful environmental and occupational epidemiologic studies
- What epidemiology is and why an interdisciplinary approach to epidemiologic study is necessary
- Frequently used epidemiologic metrics and measures of association
- Frequently used epidemiologic study designs for data collection and analysis
- How epidemiologists assess human exposure to risk factors of interest
- The use of biomarkers in environmental and occupational epidemiologic studies
- How epidemiologists infer causality

21.1 ENVIRONMENTAL AND OCCUPATIONAL EPIDEMIOLOGIC STUDIES

Associations between risk factors and human disease are exceedingly complex. Exposure scenarios are diverse. There are multifactorial etiologies and chronic low-dose risk factor exposures are usually involved. These exposures are concomitant with long latency periods to the development of many outcomes or study events of interest. It has become clear that there is a need to study intermediate causes and responses to disease. The environmental health paradigm involves mechanisms that affect the sequence of outcomes from release of an agent, or *risk factor*, involving transport, transformation, and fate processes, which lead to environmental concentrations and human exposure. There are also demographic, geographic, and lifestyle attributes that influence human exposure potential. There are toxicokinetic issues with regard to bioavailability, absorption, distribution, metabolism, and excretion that lead to internal doses and doses to the target tissues or sites. Finally, there are toxicodynamics such as compensation, damage, and repair eventually leading to the adverse outcome or effect. As noted by Saxon Graham in his seminal paper on the sociological approach to epidemiology, breaking the etiological chain anywhere along this continuum can prevent an adverse health outcome. Therefore, it is important for environmental epidemiology to explore all facets of this theoretical paradigm and most importantly to develop indicators or markers of exposure as well as disease that indicate preclinical damage. There are also susceptibility issues with much current interest in gene–environment interactions, gene–gene

Principles of Toxicology: Environmental and Industrial Applications, Third Edition. Edited by Stephen M. Roberts, Robert C. James and Phillip L. Williams.

interactions, and interactions with epigenetic modifications of the genome.

The importance of multidisciplinary efforts in environmental epidemiology was reinforced in the early 1990s. There was clear recognition of the need for cooperation among the disciplines including epidemiology and toxicology, as well as industrial hygiene and risk assessment to improve biologic monitoring of both exposures and outcomes. Several other issues that traversed many of the discussion topics were the importance of biologic mechanisms of toxicity, development of biological indicators, or biomarkers of exposure and outcomes and the value of cross-training environmental health professionals including epidemiologists and toxicologists. Madia, in an editorial in the journal *Science*, stated that environmental problems required the combined insight of a variety of scientific disciplines to understand the complexity of exposure pathways and related health effects. He went on to say that dealing with this level of complexity required not just more research, but a revolutionary new approach. Creativity and innovation in this new setting requires utilization of new technologies and the integration of various disciplines including epidemiology and toxicology.

The number of potentially hazardous agents employed in modern industry and manufacturing increases is vast; there are more than 80,000 compounds in use in the United States. Approximately 3000–4000 of these compounds are produced in high volume, and hundreds of additional chemicals are introduced into the United States on an annual basis; myriad agents are detected in human specimens. Although a majority of these compounds are likely innocuous, an integrated approach incorporating epidemiology and toxicology is increasingly required to identify and characterize human risk as illustrated in Figure 21.1.

Experimental studies are invaluable for the identification of agents posing a potential human hazard, but observational human studies are necessary to accurately quantify the concentrations of an agent that are hazardous to human health. Conversely, experimental studies are often necessary to characterize the mechanism of toxicity for an association observed in human populations, lending biologic plausibility to agent–outcome associations. Associations between potentially hazardous agents and human health outcomes are frequently identified first by occupational epidemiologists, who recognize and report on exposures in the workplace that are often higher than in the general populations and in which study populations are more well defined. Later, associations identified in the workplace by occupational epidemiologists may be considered among the general population by environmental epidemiologists. Initial concern for exposures in the workplace and/or in the environment is often prompted by the results of toxicologic investigation in the laboratory. For example, multidisciplinary evidences from toxicologic and epidemiologic studies, both occupational and environmental,

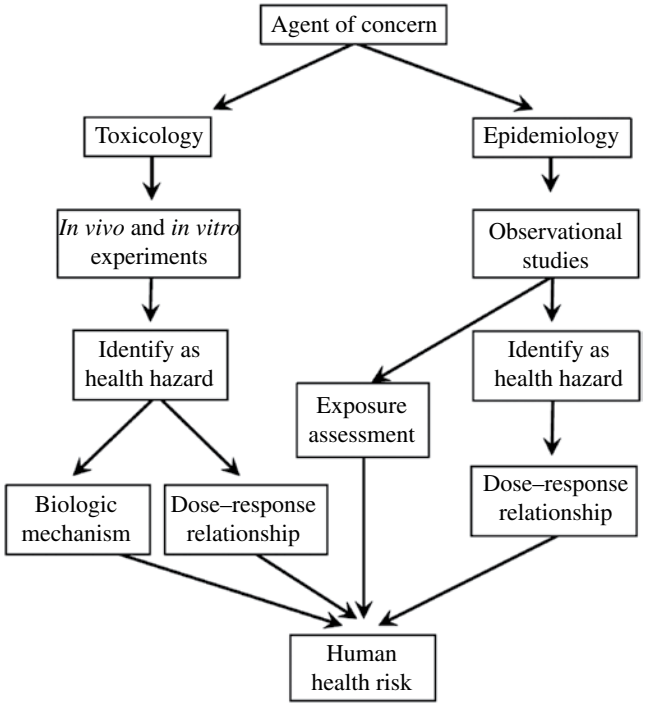


FIGURE 21.1 Integration of toxicology and epidemiology for evaluating human health risks. *Source:* Adapted from Jaffery et al. (2002).

provide the "weight of evidence" circumscribing the myriad hazards of human dioxin exposure.

Environmental and occupational epidemiology by necessity involves approaches by interdisciplinary teams that involve experts from many disciplines including exposure assessors, biostatisticians, and especially toxicologists. A prime example is the New York State Angler Cohort Study, implemented in western and central New York by Vena et al. in the early 1990s. Toxicologists from the Toxicology Research Center, at the University at Buffalo, State University of New York, and the Wadsworth Center at the New York State Department of Health were an integral part of an interdisciplinary team and developed and implemented biomarkers of exposure and intermediate toxicologic outcomes. This collaboration entailed frequent team meetings to discuss and debate analytic methods, quality assurance and control, measurement error, and consistency of measures. Dialogue concerning disciplinary methods was essential to facilitate appreciation and understanding across disciplines, to foster learning and understanding of the approaches needed for innovative and cutting-edge research, and to generate knowledge generalizable to populations of concern and that could be used for risk assessment. Therefore, toxicologists and epidemiologists must learn each other's discipline, become familiar with the special use of terms and concepts, understand these concepts, and have insights into the methodology used by each other's

discipline. A coupled role for epidemiology and toxicology is essential for discerning human health effects, and this is increasingly appreciated by experts in these respective fields. We must learn to appreciate and understand different approaches and scientific cultures and perspectives.

21.2 WHAT EPIDEMIOLOGY IS

Data collected from human studies is widely considered the "gold standard" for assessing human risk from potentially hazardous agents. However, with rare exceptions, it is impractical, infeasible, and unethical to conduct toxicologic experiments using human subjects; animal and *in vitro* models are thus invaluable tools for health risk assessment. While experimental animal and *in vitro* systems permit evaluation of an agent's toxicity under rigorously controlled conditions, extrapolation of these data to the human experience is fraught with challenges. Interspecies differences in the bioavailability of an agent due to differences in absorption, distribution, metabolism, and excretion; unrealistic exposure scenarios, genetically homogeneous test populations; and nonphysiologic conditions are just a few of the limitations complicating the use of experimental data for predicting human risk and compelling investigators to collect data among human populations.

Given the limitations inherent to the use of experimental models for human risk assessment, a nonexperimental or observational approach is required to ensure a systematic and repeatable strategy for the study of potential human health hazards. Epidemiology, translated literally from Greek roots as "…the study of that which is upon the people…" (*epi*, "upon"; *demo*, "people"; *ology*, "discourse"), is just such an approach. Using the epidemiologic approach, an investigator typically avails her- or himself of natural experiments to investigate variability in the occurrence of an outcome of interest (i.e., such as a disease) relative to a risk factor of interest (i.e., such as a chemical agent or element) and attempts to accommodate related but uninformative or nuisance sources of variability using statistical approaches. *A Dictionary of Epidemiology* defines the field as the "… study of the distribution and determinants of health-related states or outcomes in specified populations, and the application of this study to the control of health problems." The overall aim of the epidemiologic approach is to generate knowledge that can inform policies and regulations to promote, protect, and restore human health.

Environmental Epidemiology

Studies of health-related outcomes among populations with involuntary exposure to hazardous agents fall within the purview of the subdiscipline *environmental epidemiology*. Populations are generally exposed to agents in their environment through inhalation of airborne contaminants, ingestion of contaminated food or water, or dermal absorption through direct contact with soil and dust, or with consumer products. Just a few examples of studies of exposure to environmental agents and human health outcomes are investigations of arsenic consumed in contaminated drinking water sources and spontaneous pregnancy loss, dietary exposure to polychlorinated biphenyls (PCBs) through contaminated animal fats and altered thyroid hormones, absorption of high-frequency nonionizing radiation through mobile phone use and brain tumors, or investigations demonstrating associations between secondhand exposure to environmental tobacco smoke and lung cancer. Environmental epidemiology studies are challenging to conduct as these often:

- Assess risks of modest to moderate magnitude, which may be difficult to detect
- Assess very low or *trace* exposures to risk factors
- Encounter a variety of dose–response relations or thresholds for effect that may exist
- Address poorly defined study populations, which may introduce systematic errors or *bias*, and make collection of data concerning variables related to the association of interest and capture of individuals developing the outcome of interest difficult
- Address outcomes with long latency periods between exposure to a risk factor and development of a clinical outcome
- Have no group unexposed to a risk factor with which to compare exposed study participants
- Present challenges in the assessment and assignment of exposure to risk factors of interest, resulting in exposure misclassification

Despite these challenges, the results of environmental epidemiology studies are often of substantial public health relevance, as the agents examined are frequently widespread and thus even small increases in risk translate to large numbers of *attributed* outcomes.

Occupational Epidemiology

Studies of health-related outcomes among populations with work-related exposure to potentially hazardous agents fall within the purview of the subdiscipline *occupational epidemiology*. For example, this might include investigations of associations between occupational diesel exhaust inhalation and lung cancer, exposure to toluene isocyanate during the manufacture of products from polyurethane foam and asthma, Percival Pott's classic eighteenth-century study of scrotal cancer among young English chimney sweeps, or Irving Selikoff's studies of mortality and asbestos exposure among insulation manufactures in the mid-twentieth century.

Epidemiologic studies in occupational settings offer several advantages compared to environmental epidemiology studies, including:

- Assessing risks of moderate to large magnitude, which are often easier to detect
- Having fairly high or consistent exposures to risk factors of interest
- Addressing well-defined study populations, which minimizes the introduction of systematic errors, and facilitating capture of data concerning variables related to the association of interest as well as capture of individuals developing the outcome of interest
- Having a group unexposed to the risk factor of interest with which to compare exposed study participants
- Having well-described exposures to risk factors of interest associated with job duties and responsibilities
- Having employment records useful for determining exposure to risk factors and capturing study outcomes

Despite some of the advantages offered by occupational health investigations, the results of such studies are also limited by several issues. The underlying distribution of risk for an outcome in a working population is likely to differ from that in the general population; employed persons are healthier on average than the unemployed component of the general population. This *healthy worker effect* may obscure associations between risk factors and outcomes of interest under certain conditions. Although exposure assessment is often more tenable in occupational epidemiology studies, as compared to environmental epidemiology studies, accurately assessing personal exposure to agents also presents with challenges in the occupational setting. Nonetheless, occupational epidemiologic studies often offer the opportunity to study high-dose exposures that can be extrapolated to the lower doses seen in environmental studies. Almost inevitably, the agents used in the occupational setting are discharged or released into the environment resulting in pathways to human exposure in far larger populations.

21.3 EPIDEMIOLOGIC METRICS AND MEASURES OF ASSOCIATION

To detect and quantify associations between risk factors and outcomes, or events of interest, epidemiologists employ a variety of metrics including rates and ratios, with the ultimate aim of assessing evidence for a causal relation. A selection of frequently used epidemiologic metrics can be found in Table 21.1. Epidemiologists count the number of outcomes of interest including deaths, diseases, injuries, or diagnoses, to name but a few, and describe these in terms of *person* (who is affected and who is not affected?), *place* (where are people affected or not affected?), and *time* (when are people affected or not affected?). These *descriptive epidemiology* data are useful for generating *hypotheses*, educated testable conjectures describing relations between environmental or occupational exposures to risk factors and study outcomes. Two types of study outcomes are counted, new or *incident* outcomes and existing or *prevalent* outcomes. To facilitate comparison between groups of different sizes and with different attributes, the frequency of an outcome is often expressed in terms of a risk or a rate—the number of outcomes experienced by a population unit or population–time unit across or during a time interval.

Incidence

The *incidence* describes the probability for a new outcome in a specified population at risk—the number of new outcomes that accrue in a population at risk (Table 21.1). The incidence may be expressed in terms of numbers of people at risk for an outcome (i.e., *cumulative incidence proportion*) or in terms of the person-time at risk for an outcome (i.e., *incidence density rate*). The incidence is interpreted as the *risk* for an outcome. The instantaneous incidence rate, meaning the incidence rate over a time interval approaching zero, is referred to as the *hazard rate* and is a quantity conditioned on the probability that a participant did not experience an earlier outcome. These types of incident measures are essential for establishing causal associations between a risk factor and outcome of interest. For example, after accommodating difference in age distributions (i.e., *age adjusted* or *age standardized*), the U.S. incidence proportions for breast cancer were 121.0 cases per 100,000 white women and 117.0 cases per 100,000 black women in 2007. Thus, 121 new cases of breast cancer were diagnosed for each 100,000 White women, and 117 new cases of breast cancer were diagnosed for each 100,000 Black women in the U.S. population during 2007.

Prevalence

Somewhat different from the incidence rate, the *prevalence proportion* describes the probability of an existing outcome in a population at a point in time (i.e., *point prevalence*) or during a time interval (i.e., *period prevalence*); prevalence includes past as well as new outcomes (Table 21.1). In a *stationary population* (i.e., no migration), the prevalence proportion is a function of the incidence rate and the duration of the study outcome such that prevalence $\approx$ incidence $\times$ duration. The prevalence proportion, while useful, cannot be interpreted as a risk as the duration of the outcome is presumed to vary by person, and thus, variability in prevalence will in part reflect time to recovery or death, contingent on the outcome. Somewhat rare outcomes with long duration, such as infertility (i.e., possibly lasting for many

TABLE 21.1 Common Epidemiologic Metrics and Measures of Association

Metric	Use	Formula
Cumulative incidence proportion	Describe the risk for an outcome across a time interval	$\left[\frac{\#\text{ New outcomes}}{\text{Population size at risk during a time interval}}\right]\times 10^n$ persons
Incidence density rate	Describe the rate for an outcome during a time interval	$\left[\frac{\#\text{ New outcomes}}{\text{Person-time at risk during a time interval}}\right]\times 10^n$ person-years
Point prevalence proportion	Describe the number of existing outcomes (old and new) at a single time	$\left[\frac{\#\text{ Existing outcomes}}{\text{Population size}}\text{ at a point in time}\right]\times 10^n$ persons
Period prevalence proportion	Describe the number of existing outcomes (old and new) across a time interval	$\left[\frac{\#\text{ Existing outcomes}}{\text{Population size}}\text{ over a time interval}\right]\times 10^n$ persons
Relative risk	Describe the relative cumulative risk for an outcome across a time interval between exposed and unexposed groups	$\frac{\text{Cumulative incidence proportion}_{\text{Exposed to a risk factor}}}{\text{Cumulative incidence proportion}_{\text{Unexposed to a risk factor}}}$
Rate ratio	Describe the ratio of outcome rates during a time interval between exposed and unexposed groups	$\frac{\text{Incidence density rate}_{\text{Exposed to a risk factor}}}{\text{Incidence density rate}_{\text{Unexposed to a risk factor}}}$
Hazard ratio	Describe the relative instantaneous risk for an outcome between exposed and unexposed groups	$\frac{\text{Hazard rate}_{\text{Exposed to a risk factor}}}{\text{Hazard rate}_{\text{Unexposed to a risk factor}}}$
Prevalence proportion ratio	Describe the relative prevalence proportion for an outcome between exposed and unexposed groups	$\frac{\text{Prevalence proportion}_{\text{Exposed to a risk factor}}}{\text{Prevalence proportion}_{\text{Unexposed to a risk factor}}}$
Attributable risk	Describe the additional number of outcomes associated with exposure to a risk factor	$\text{Incidence rate}_{\text{Exposed to a risk factor}}-\text{Incidence rate}_{\text{Unxposed to a risk factor}}$
Attributable fraction	Describes the proportion of outcomes in the exposed group, associated with exposure to a risk factor	$\frac{\text{Incidence rate}_{\text{Exposed to a risk factor}}-\text{Incidence rate}_{\text{Unexposed to a risk factor}}}{\text{Incidence rate}_{\text{Exposed to a risk factor}}}\times 100$
Odds	Estimate the odds for an outcome relative to no outcome	$\frac{\text{Probability for outcome}}{1-(\text{Probability for outcome})}$
Odds ratio	Estimate the underlying population relative risk using the ratio of odds for an outcome between exposed and unexposed groups	$\frac{\text{Odds}_{\text{Exposed to a risk factor}}}{\text{Odds}_{\text{Unexposed to a risk factor}}}$

years) may have a high prevalence rate that belies a low incidence rate. In contrast, fairly common outcomes of short duration such as influenza (i.e., resolves or results in death in short time) may have a high incidence rate that belies the comparatively low prevalence rate. As a reflection of existing cases of a disease, injury, or condition, prevalence data are useful for formulating public policy and budgeting resources, as well as for generating hypotheses. For example, the prevalence of major congenital malformations has been reported to be 3% in U.S. live births; this value indicates that 3 of every 100 live infants have a major congenital malformation and reflects both the incidence of the malformation during gestation and the survival of a malformed fetus to delivery. Given that collection of incidence data is often costly and time-consuming as a population must be followed over time for the occurrence of new outcomes, prevalence data are often more readily available or readily collected for epidemiologic study.

Ratios

Although incidence and prevalence measures are of great utility in characterizing outcomes in terms of person, place, and time, these alone provide little insight into the causes of those outcomes. However, ratios of these values for different groups can provide valuable information in terms of the relative difference in incidence or prevalence. The ratio of incidence proportions or rates for an outcome among

persons exposed to a risk factor of interest, such as a specific chemical agent, relative to the outcome among persons unexposed to that risk factor (i.e., *referent* group) is referred to as the *relative risk* (RR) or *rate ratio* (Table 21.1). Various ratios, including ratios of risks, rates, hazards, and prevalence proportions, are used by epidemiologists to characterize associations between risk factors and outcomes of interest. Ratios are unitless quantities with values exceeding one suggesting a higher probability for the outcome among persons exposed to a risk factor (i.e., *adverse effect*), values less than one suggesting a reduced probability for an outcome among persons with a risk factor (i.e., *protective effect*), and a value equal to one suggesting no association between a risk factor and an outcome. For example, in a hypothetical study of the association between female blood Cd level and infertility among couples trying to conceive (i.e., no pregnancy), an RR equal to 1.82 would indicate an 82% increase in the probability for infertility among women with high blood Cd concentrations, compared to women with low blood Cd concentrations.

Attributable Risks

The *attributable* or *excess risk* describes the additional number of outcomes, or events of interest, experienced in association with exposure to a risk factor. It is defined as the absolute difference in incidence measures for the outcome of interest between those with exposure and those without exposure to the risk factor of interest (Table 21.1). One common interpretation of this value is that it represents the number of outcomes that would be eliminated in exposed members of the study population should the risk factor of interest be removed. For the aforementioned hypothetical example, a cumulative incidence rate of 182 cases of infertility (i.e., no pregnancy) per 1000 women with high blood Cd and 100 cases of infertility per 1000 women with low blood Cd suggests that 82 women did not achieve pregnancy due at least in part to high blood Cd.

By extension, the *attributable fraction* describes the proportion of cases experienced by the exposed group, in association with the risk factor of interest. Although there are several approaches to operationalize this quantity, one common definition entails division of the attributable risk by the incidence measure for an outcome among those exposed to the risk factor of interest, and then the quotient is multiplied by a constant (Table 21.1). A frequent interpretation of this value is that it represents the proportion of outcomes, among the exposed, that would be eliminated from the study population should the risk factor of interest be removed. For the earlier hypothetical example, the attributable fraction for Cd and pregnancies suggests that decreasing Cd exposure would decrease the number of infertile women with high Cd by 45%.

Odds

Under some circumstances, such as the methodologic constraints introduced by the case–control design (to be described later in this chapter) or to simplify a statistical analysis, the *odds* of a study outcome may be estimated. Odds are not a proportion or a rate, but rather represents the ratio for the probability that an outcome occurred to the probability that the outcome did not occur (Table 21.1). In the earlier hypothetical example the odds for infertility among women with high blood Cd is 0.22, whereas the analogous odds for women with low blood Cd is 0.11.

The *odds ratio* characterizes the odds for an outcome among the group exposed to a risk factor of interest relative to the odds for the outcome among a group unexposed. Odds ratios are interpreted in a fashion analogous to rate ratios, with their meaning contingent on the design of the epidemiologic study from which the data emerge (i.e., there are prevalence odds ratios, incidence odds ratios, and exposure odds ratios). Under a limited set of circumstances, principally that the outcome under study is rare in the study population from which the study sample was recruited (i.e., <10%), the odds ratio for an outcome reflects the RR for that outcome. However, with increasing prevalence of the outcome in the study population, the odds ratio increasingly overestimates the RR. For example, in the hypothetical study described earlier, the odds ratio for infertility among women with high blood Cd is 2.0, which modestly overestimates the aforementioned RR equal to 1.82. However, the prevalence of infertility is 14.1 cases per 100 women in the hypothetical study population (i.e., 14%).

Random Error in Epidemiologic Studies

Given that complete ascertainment of a study population or *census* is frequently infeasible and impractical, epidemiologists employ sampling in which a smaller group from the population is studied, and the results are then generalized to the greater whole. Ideally, an investigator aspires to recruit a *study sample* that reflects the distribution of risk factors, study outcomes, and other important factors in the source *study population*; however, this can be quite challenging. Failure to recruit a sample representative of the study population limits the *generalizability* of study results; the investigator is unable to extrapolate the results to the study population leading to limited *external validity*.

For example, in Figure 21.2, three study samples of $n = 200$ are taken from a hypothetical study population of 1000 people. Study sample "a" reflects the population in terms of appearance and age distribution; there is no sampling error. However, study samples "b" and "c" appear different from the study population as reflected in their disparate age structures; there is substantial *sampling error*. When a sampling error occurs in a *nondifferential* fashion, meaning it does not vary by the risk factor or outcome of interest or by

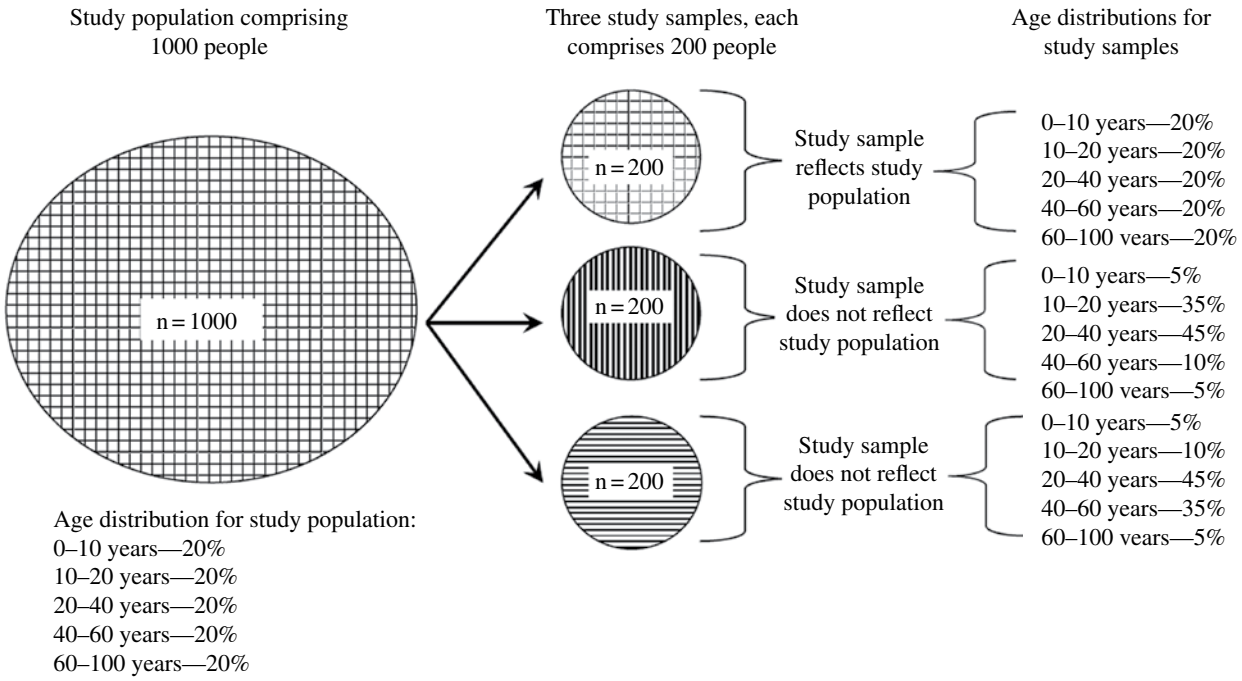


FIGURE 21.2 Selection of three study samples from a hypothetical population of 1000 persons.

any other important study variable, it is defined as *random*. Should sampling error occurs in a *differential* fashion, meaning it varies by the risk factor of interest, outcome of interest, or another important study factor, it is referred to as *biased*; this type of systematic error undermines the *internal validity* of a study (i.e., the results generated in the study sample). Internal validity is of primary concern as this is a prerequisite to external validity; vis-à-vis, the results in the study sample must first be valid to extrapolate these to the study population. While preventable through appropriate study techniques, bias is insidious and difficult to assess and quantify, and once introduced, it is difficult if not impossible to eliminate. Epidemiologic bias will be discussed in greater detail later in this chapter.

All study results are associated with error, including incidence rates, rate ratios, and other epidemiologic metrics. Epidemiologists will frequently assess study results for *statistical significance* in one fashion or the other; inherently, this quantifies the role that random error (i.e., *chance*) plays in a study result. Under limited conditions, hypothesis testing offers the investigator guidance as to the role played by random sampling error in the observed study results. The *type 1 error*, or "α," describes the probability for detecting an association in the study sample that does not exist in the study population (i.e., a *false-positive* result due to sampling error). By convention, type 1 error is often set at a threshold of 5%, meaning that a result likely to occur randomly in less than 1 of 20 studies is considered valid. However, many other approaches to quantifying random error are also employed. Often, epidemiologists quantify the random error associated with a study result using *P-values* (i.e., using thresholds such as $p<0.05$) or *confidence intervals* (i.e., such as 95% confidence intervals).

A study sample must also be of sufficient size, given the hypothesis, to find an association if one does indeed exist in the study population. This potential is referred to as statistical *power*, assessed as $1-\beta$. The quantity "β" is referred to as the *type 2 error*, and it describes the probability for not detecting an association in the study sample, which does exist in the study population (i.e., a *false-negative* result due to sampling error). Type 1 and type 2 errors are inversely related in that increases in the type 1 error (i.e., corresponding to decreasing *specificity*) are associated with decreases in the type 2 error (i.e., corresponding to greater *sensitivity*), and vice versa. By convention, at least 80% power is required for a study to be considered sufficiently precise, as power is inversely associated with random error (i.e., greater statistical noise inhibits an investigator's ability to detect a true signal). Thus, it behooves an investigator to reduce random error whenever possible. However, studies are frequently conducted with power less than 80%, substantially so in many cases. For a thorough treatment of these issues, the reader is referred to the additional sources listed at the end of this chapter.

21.4 EPIDEMIOLOGIC STUDY DESIGNS

The design of observational studies accounts for much of the discipline of epidemiology, with specific adaptations particular to various subdisciplines including environmental

and occupational epidemiology. An epidemiologic study generally proceeds in three stages: (i) the *design stage*, (ii) the *implementation stage*, and (iii) the *analysis and reporting stage*. It is important for toxicologists to be involved in every stage of an epidemiologic study to provide the insight and expertise from that discipline. In the design stage, a study hypothesis is refined and developed as the specificity and nature of the question should govern the remainder of the design process; the hypothesis serves as the anchor for an epidemiologic study. In the design stage, a study population is identified and a strategy to recruit a representative and unbiased study sample is prepared; this step will facilitate the success or failure of an investigation. In the implementation stage of an epidemiologic study, participants are recruited and data and possibly biologic specimens are collected and analyzed; this phase may be of limited duration such as under a *cross-sectional study design*, or it may take many years such as under a *prospective cohor*t study design. *Randomized controlled trials*, though epidemiologic in nature, are rarely encountered in the investigation of exposure to environmental and occupational risk factors. This quasiexperimental study design is frequently used to investigate pharmaceutical treatments, medical procedures, medical devices, or other health interventions and so is not addressed in this chapter. The features, strengths, and limitations of the most commonly employed study designs in environmental and occupational epidemiology are summarized in Table 21.2.

Many of the obstacles to successful study implementation can be prevented by meticulous planning during the design stage. The decisions made in the design stage of a study are frequently indelible once a study moves into the implementation stage. Biases introduced deliberately or inadvertently may be impossible to eliminate during the analysis phase of the study, if not prevented or if appropriate and sufficient data are not collected. There are of course some issues to which the investigators may adapt during the course of implementation. However, these may present additional challenges, as the introduction of differential participant recruitment strategies or data collection methodologies during the implementation stage may invalidate previously recruited participants or previously collected data. Very careful attention must be paid by the investigators to the myriad nuances of subject recruitment and data collection a priori. In the analysis and reporting stage, data are *cleaned*, statistical analysis is conducted, and the results are reported in oral and written modalities as conference presentations, peer-reviewed publications, and technical reports. During the analysis and reporting stage, the investigator will have the opportunity to adjust or control for certain biases due to confounding contingent on collection of the requisite data (discussed in detail later in this chapter), to evaluate and on occasion adjust or control other sources of bias, and even to conduct additional hypothesis testing in the so-called secondary studies.

Here, we briefly summarize the most salient components, limitations, and strengths of the most widely used epidemiologic study designs. The reader is referred to the sources listed at the end of this chapter for a more comprehensive treatment. Epidemiologic studies are traditionally dichotomized as *descriptive* or *analytic/inferential*. In descriptive studies, investigators seek to characterize an outcome of interest in terms of factors related to person, place, and time as described earlier in this chapter. These studies generally provide data concerning the frequency and rates of outcomes and exposure to risk factors of interest; these data are invaluable for hypothesis generation as well as for surveillance, policy development, and program planning and evaluation. *Case studies*, *case series*, and *surveillance studies* are types of descriptive studies. Analytic/inferential studies focus on statistical tests of etiologic hypotheses with the ultimate goal of inferring causal associations. Analytic/inferential epidemiologic study designs are defined in terms of the strategy employed to recruit study participants and the *unit of analysis* used. Most analytic studies include a descriptive component. Two general analytic/inferential strategies are implemented as defined by the analysis at the *group level* or the *individual level*.

Group-Level Epidemiologic Study Designs: Ecologic Studies

In group-level or *ecological* study designs, data are averaged over groups of individuals that serve as the unit of analysis. Groups are often fixed at a single point in time and vary geographically, or alternately groups may be fixed geographically and vary temporally. Under a traditional *spatial ecologic* design, the incidence or prevalence of an outcome is compared across different geographic units that have different distributions for a risk factor of interest; attempts are frequently made to ensure that groups are otherwise similar but this may prove challenging. For example, studies report higher rates of lung cancer in regions with very high concentrations of inorganic As contaminating ground drinking water sources than in regions without such contamination.

Ecologic studies are often feasible using existing data sources, so these may be more convenient and less expensive than study designs requiring individual-level data, which is typically more resource intensive. However, the results generated by ecologic studies are highly limited and are considered useful primarily for hypothesis generation but not for causal inference. Foremost among the limitations of the ecologic approach is the *ecologic fallacy*; associations between groups may differ substantially from associations between individuals due to differential heterogeneity of within-group study factors. Furthermore, under the spatial ecologic approach, *temporality*, in which exposure to a risk factor chronologically precedes the outcome of interest, may not be established, thereby introducing possible *reverse*

TABLE 21.2 Common Epidemiologic Study Designs

Design	Basic Features	Strengths	Limitations
Descriptive studies	Count outcomes in terms of person, place, and time	Provide data for public health policy development, program planning, and budget allocation; useful for generating hypotheses for future study	Unable to assess associations between risk factors and outcomes or to provide information to infer causality
Ecologic studies	Compare rates for outcomes at the group level of measurement; a group exposed to a risk factor is compared to a group unexposed to a risk factor; groups are defined in terms of geographic space or time	Usually less resource intensive than other designs, often can make use of readily available data collected as part of descriptive studies, useful to infer causality at the group level, useful for hypothesis generation at the individual level	Results valid only at the group level of measurement (vulnerable to the ecologic fallacy), difficult to address confounding at the individual level of measurement, unable to assess temporality in some approaches and thus vulnerable to reverse causation
Cross-sectional	A "snapshot" in time at the individual level of measurement; exposure to a risk factor is assessed simultaneously with outcomes of interest	Usually less resource intensive than other study designs, uses individual level data, and permits simultaneous assessment of multiple risk factors and outcomes; useful for hypothesis generation	Captures prevalent outcomes and so cannot directly estimate risk, unable to assess temporality and thus vulnerable to reverse causation, vulnerable to recall bias
Case–control	Compares individuals with a study outcome (cases) to those without (controls) and retrospectively assesses exposure to risk factors	Very useful for rare study outcomes; permits simultaneous assessment of multiple risk factors, uses individual-level data, and incorporates temporality between risk factor and study outcome; often less resource intensive than cohort studies	Usually limited to a single study outcome; control participants can be difficult to identify and recruit; vulnerable to selection bias and recall bias; retrospective exposure assessment is vulnerable to exposure measurement misclassification bias; estimates population relative risk only when the study outcome is rare or using a more complex incidence density sampling strategy
Cohort	Among individuals free from a study outcome, compares a group exposed to a risk factor of interest to a group unexposed to a risk factor of interest and follows through time for occurrence of study outcomes	Provides risk estimates and unbiased estimates for population ratios, permits simultaneous assessment of multiple study outcomes, and incorporates temporality between a risk factor and study outcome; prospective exposure assessment	Usually limited to one or a few risk factors of interest; study dropouts (loss to follow-up) may introduce selection bias; rare study outcomes may require very large study samples to facilitate statistical analysis; study outcomes with long latency periods may require long periods of follow-up; often highly resource intensive and time-consuming

causality, in which the outcome of interest alters the level of exposure to the risk factor. In addition, adjustment for confounding and consideration of effect measure modification (issues discussed in detail later in this chapter) may be more challenging than when using individual-level data study designs; group-level data for these factors is also subject to the aforementioned ecologic fallacy. Thus, interpretation of study results made at the group level may have no meaning or may have an entirely different meaning at the individual level. However, circumstances in which factors exclusive to groups, such as implementation of a new public policy in a specific legislative district, or in which extrapolation to groups are of primary interest may be uniquely apropos to the spatial ecologic study design.

A design related to yet distinct from the spatial ecologic design is the *time-series* design. This framework is essentially an ecologic design in which risk factors and outcomes are grouped by time interval and are often fixed in space. This approach is very useful to investigate the effects of environmental or occupational exposures that change over time in a single geographic area. Like the traditional ecologic design, time-series studies are vulnerable to ecologic bias and confounding bias. In addition, time-series designs present with several unique challenges including *autocorrelation*, in which observations across time are associated with one another (i.e., most statistical analyses presume *independent* study outcomes); *latency*, in which an outcome associated with exposure to a risk factor during an earlier time interval may manifest during a later time interval; and bias due to time-varying confounders; fairly complex statistical approaches may be required to accommodate these issues. Time-series study designs have been used to great effect for investigating the effects of air pollution on human health outcomes. For example, a substantial increase in mortality rates in London, England, primarily associated with respiratory disease, was documented shortly after the infamous London Fog pollution "Pea Soup" episode of 1952. In that scenario, widespread use of coal for household heating coincided with a temperature inversion trapping airborne pollutants close to the earth's surface and resulting in widespread respiratory exposure to sulfur oxides and particulates over a limited duration.

Individual-Level Epidemiologic Study Designs: Cross-Sectional Studies

At the individual level, data are collected from each study participant, and thus, study designs are not vulnerable to the aforementioned ecologic fallacy. There are several individual-level study designs commonly implemented by epidemiologists, and there are many variations on each theme. The simplest, quickest, and usually least resource-intensive individual-level approach is the *cross-sectional study design*. The cross-sectional design essentially comprises a "snapshot"

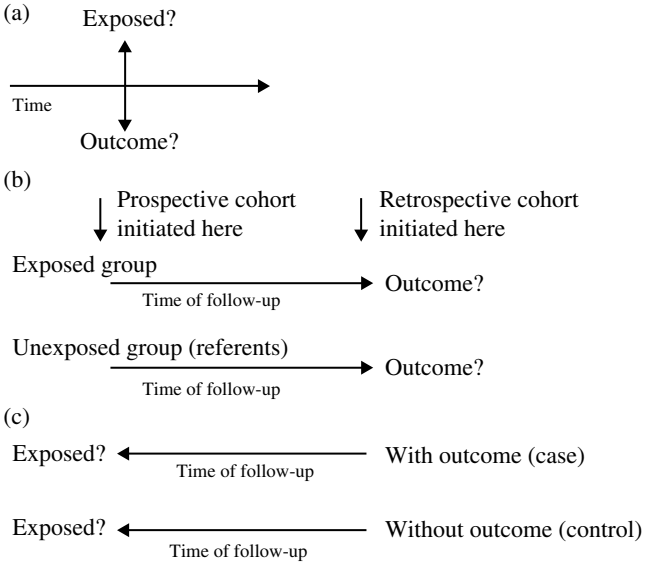


FIGURE 21.3 Common epidemiologic study designs: (a) cross-sectional design, (b) cohort design, and (c) case–control design.

in time as described by Figure 21.3a. Participants are recruited to the study sample from a study population at a single instance or over a single time interval. Exposure to one or more risk factors of interest and one or more outcomes of interest are assessed simultaneously. Thus, participants are selected without regard to exposure or outcome, and thus, the study sample provides an estimate of prevalence for both in the study population. This study design is among the most frequently found in the environmental and occupational epidemiology literature.

It is important to note that the cross-sectional approach captures only prevalent cases; it is also referred to as a *prevalence study*. As a consequence, associations between exposure to risk factors and outcomes may be a consequence of association with the duration of the outcome rather than with the incidence of the outcome, as discussed earlier in this chapter. Moreover, this approach precludes assessment of temporality, and thus, the cross-sectional design is vulnerable to the earlier discussed reverse causality. Cross-sectional studies provide prevalence rate ratios or odds ratios to characterize associations between one or more risk factors and one or more study outcomes. For example, an investigation that simultaneously measures PCBs, dioxins, and thyroid hormones in the blood of study participants, at a single point in time, and then evaluates these data for statistical associations between the PCBs and dioxins with thyroid hormones is cross-sectional in nature. Not infrequently, an investigator may target recruitment of subjects likely to have higher or lower exposures or subjects more likely or less likely to have experienced an outcome of interest in order to increase the available statistical power (i.e., greater variability in exposure to risk factors generally increases statistical power). In the aforementioned cross-sectional investigation of PCBs and

TABLE 21.3 Distribution of Participants in Epidemiologic Studies: Cross-Sectional (Prevalence Data) or Cohort (Incidence Data) Study Designs

	Outcome (Cases)	No Outcome	Total
Exposed to risk factor	A	B	A+B
Unexposed to risk factor (referents)	C	D	C+D
	A+C	B+D	A+B+C+D

Outcome proportion or risk in the exposed, A/(A+B); outcome proportion or risk in the unexposed, C/(C+D); prevalence proportion ratio or relative risk, A/(A+B)/C/(C+D)

dioxins and thyroid function, an investigator might recruit individuals likely to have high and low rates of Great Lakes sport fish consumption, such as licensed anglers living in close proximity to the Lake Ontario, to ensure a wide distribution of exposure to the risk factors (lipophilic compounds including PCBs and dioxins are known to concentrate in the fat of Great Lakes sport fish).

Table 21.3 describes a generic cross-sectional study of a single risk factor and study outcome of interest using a 2×2 table, or *contingency table*. Study participants are classified according to exposure status and then distributed across columns according to whether or not the outcome of interest occurred. These data can be employed to calculate prevalence proportions or odds and ratios described in Table 21.1.

Individual-Level Epidemiologic Study Designs: Cohort Studies

Individual-level epidemiologic studies more complex than the cross-sectional design incorporate temporality, although these usually require a greater investment in time and resources than the cross-sectional approach. In a *cohort study*, a sample of persons not having experienced the outcome of interest (i.e., a "cohort") is followed over time for occurrence of that outcome as described in Figure 21.3b. A group of participants exposed to the risk factor of interest is followed as is a comparable group that is unexposed to the risk factor of interest, a referent group. At the end of a defined study period, or period of *follow-up*, incidence measures for outcomes in the exposed group are compared to the unexposed group. A cohort comprised of participants not having experienced the study outcome and exposed to a risk factor, as well as participants not having experienced the study outcome and unexposed to a risk factor, can be identified at baseline and then followed *prospectively* through time until occurrence of the study outcome or until the end of the follow-up period. This approach can often be enormously time-consuming contingent on the latency for the outcome of interest, such as a solid tumor that may take 20 years or longer to develop, and thus can be prohibitively expensive. Alternately, persons may be identified as exposed or unexposed to a risk factor of interest *retrospectively*, using preexisting records or archived biologic specimens collected in the past, and figuratively followed up for the study outcome of interest or the end of the study period, having occurred in the past or occurring in the present. Occupational or medical insurance records can often facilitate this process assuming the availability of accurate and unbiased medical records, and this approach is frequently used for occupational epidemiologic investigations.

Unlike the aforementioned cross-sectional design, the cohort design can provide valid risk estimates as only incident outcomes are captured; the incidence in the referent group (unexposed) characterizes the *background risk* for an outcome in the study population. These data are then used to generate RR, risk or rate ratios, or attributable risks to characterize the association between a risk factor and one or more study outcomes. The cohort study design is very useful when investigating exposure to rare or infrequently encountered risk factors as exposed individuals are targeted for recruitment. However, this approach also limits the number of risk factors that can be studied, as these must be decided a priori and a group recruited for each risk factor of interest, or combination thereof. In contrast, the cohort study affords an investigator the opportunity to assess multiple study outcomes contingent on the size of the study sample; rare outcomes may not manifest in sufficient number to facilitate statistical analysis unless the cohort is unusually large. The latter caveat makes the cohort design less useful for the study of rare outcomes.

Under some circumstances, one or more participants may leave or drop out of a cohort study prior to experiencing a study outcome and before the end of the study, and so follow-up must be terminated or *censored* at that point; this is referred to as *right censored*. Conversely, follow-up for one or more participants may be initiated after a study has begun, and thus, follow-up prior to enrollment in the cohort must be censored for a subject at that point; this is referred to as *left censored*. When censoring of follow-up takes place, there will be heterogeneity in the time individuals in the study sample are at risk for an outcome. In this scenario, epidemiologists often employ *survival analysis* approaches to analyze data, as these utilize conditional probabilities to accommodate the variability in time at risk for individuals in the study. The reader is referred to the sources listed at the end of this chapter for additional information. If censoring is related to the study outcome of interest, then a *selection bias* may be introduced into a cohort study (discussed later in this

chapter). As with all epidemiologic study designs, cohort studies are susceptible to confounding bias (discussed later in this chapter).

Cohort studies are used to generate various ratios, which as discussed are used to evaluate the effect of a risk factor on an outcome of interest. For example, the Harvard Six Cities Study, a prospective cohort investigation of air pollution and mortality, detected increased risks for respiratory and cardiovascular-related mortality in association with long-term exposure to higher ambient concentrations of fine airborne particulates. Table 21.3 describes a generic cohort study of a single risk factor and study outcome of interest using a contingency table. Study participants are classified according to exposure status, with row margins fixed by the investigator (i.e., the number of participants exposed to the risk factor and the number of participants unexposed to the risk factor, or referents), and then distributed across columns according to whether or not the outcome of interest occurred. These data can be employed to calculate outcome risks and subsequently the RR as described in Table 21.1.

Individual-Level Epidemiologic Study Designs: Case–Control Studies

In contrast to the epidemiologic cohort study design, which may be implemented in a prospective or retrospective fashion, the *case–control study design* is strictly retrospective (it is not infrequently referred to as a r*etrospective study design*). The apparent simplicity of the case–control design belies its true complexity, a comprehensive treatment of which is beyond the scope of the current chapter but which is available in the sources listed at the end of this chapter. The case–control study design is quintessentially epidemiologic; it essentially entails strategies for sampling from an underlying cohort so that associations between risk factors and outcomes can be assessed in a less resource and time-intensive manner than that offered by cohort designs. As described by Figure 21.3c, subjects experiencing the outcome of interest in the study population are recruited as *cases* (all cases in a study population or a random sample thereof), and a sample of individuals not experiencing the outcome at the time of the study are recruited as *controls*. It is important to note that controls are not necessarily healthy individuals, but are those who have not experienced the outcome or the study event of interest. This study design is very useful for rare outcomes, such as cancers or congenital malformations, which would require a very large cohort to capture a sufficient number of incident outcomes to facilitate statistical analysis. Under the case–control design, the investigator creates in essence a sample with a high proportion of cases irrespective of presence in the study population; thus, the prevalence in the study population cannot be estimated from the study sample. The investigator will choose the number of cases to include as well as one or more controls per case for purposes of comparison. To be valid, recruitment of case and control participants must be independent of their exposure to the risk factors of interest, and controls must represent the exposure experience in the study population (i.e., if the controls had indeed been cases, they would have been captured as such by the study). Otherwise, a *selection bias* may be introduced into the study. Contingent on the outcome and risk factor of interest, case–control studies are also vulnerable to *recall bias* and *exposure misclassification bias* (described later in this chapter).

As described by Figure 21.3c, case and control participants are each retrospectively assessed for exposure to a risk factor or risk factors of interest, often through interviews, archived biologic specimens, or previously collected records. Study participants may be recruited using a *cumulative incidence sampling* strategy in which cases and controls are identified at the end of a defined study period. Under this scenario, a single subject can participate in the study only as a case or as a control. Alternately, participants may be recruited using an *incidence density sampling* strategy, in which a cohort is followed and incident cases are recruited to the study sample as these occur; controls are then selected from among those individuals in the study population without the outcome at the time of each case. Perhaps counterintuitively, an individual may participate in such a case–control study as both a control and a case. However, statistical analysis of data arising from this type of study requires a specialized approach.

Variations on the case–control approach include the *nested case–control study design* and the *case-crossover study design*. In a nested case–control study, cases and controls are sampled from the participants in a completed or partially completed cohort study, providing the advantage of complete enumeration (i.e., outcomes are known for all members of the cohort). In some scenarios, costs may be prohibitive, so that only a sample of the cohort can be assessed for an expensive laboratory analysis or for a secondary hypothesis. Nested case–control study participants are "sampled" from the cohort study, which then serves as the study population of interest. In the case-crossover approach, participants with outcomes are identified, and then various *reference intervals* are sampled for each participant, so that a participant serves as her or his own control. This design presents the advantage of eliminating confounding bias by non-time-varying factors, such as genetic polymorphisms and sex. The case-crossover study design is usually employed to assess the acute effects of transient risk factor exposures; it is frequently employed to evaluate acute human health effects associated with exposure to air pollutants.

Case–control studies are used to generate odds ratios, which as discussed earlier provide an unbiased estimate of the underlying study population RR if the outcome is rare. For example, a series of case–control studies conducted by

TABLE 21.4 Distribution of Participants in Epidemiologic Studies: Case–Control Study Design

	Exposed to Risk Factor	Unexposed to Risk Factor	Total
Outcome (cases)	A	B	A+B
No outcome (controls)	C	D	C+D
	A+C	B+D	A+B+C+D

Exposure odds in cases, A/B; exposure odds in controls, C/D; odds ratio, AD/BC.

the U.S. National Institute of Occupational Safety and Health (NIOSH) established a link between occupational exposure to vinyl chloride gas and hepatic angiosarcoma, a rare form of liver cancer. Table 21.4 describes a generic case–control study of a single study outcome and risk factor of interest using a contingency table. Participants are classified according to case or control status, with row margins fixed, and then distributed across columns according to the presence (i.e., exposed) or absence (i.e., unexposed) of the risk factor of interest. These data can be employed to calculate exposure odds and subsequently the odds ratio described in Table 21.1.

Bias in Epidemiologic Studies

Study bias is defined as a directional deviation of study sample results from the true underlying study population value; vis-à-vis, effect estimates systematically shifted too high or too low. Bias in epidemiologic studies often results from unintentional trends in the collection, analysis, interpretation, publication, or review of data. Bias undermines the aforementioned validity of a study and contingent on manifestation may compromise the earlier described internal and/or external study validities. There are essentially three sources of bias in epidemiologic studies, the manifestations of which are often insidious and frequently specific to a particular study design, including (i) *information bias*, (ii) *selection bias*, and (iii) *confounding bias*.

Contingent on study design, information bias and selection bias are often preventable. However, once introduced, an investigator is usually unable to exclude these biases and is limited to characterizing the impact on study results, sometimes using *sensitivity analysis* or using external information or data to compensate for a known bias. Sensitivity analysis comprises evaluation of changes in study results when subjects, variables, or approaches are excluded or changed to assess quantitatively the effects of those components. In contrast, it is often feasible to correct for confounding bias by statistical adjustment or exclusion. This is subject to the collection of data describing anticipated confounders during the design and implementation phases of the study. Dialogue on these issues between the epidemiologist and toxicologist will determine the quality of the investigation.

Information bias results from faulty assessment of exposure or outcome in a fashion that is differential between groups compared; information accuracy differs between exposed and unexposed persons, between cases and noncases, or between both. Types of information bias commonly encountered by epidemiologists include *recall bias*, *interviewer bias*, and *misclassification bias*. Selection bias is produced when study recruitment or participation occur contingent on exposure to the risk factor under the case–control study design, contingent on the outcome under the cohort study design, or in which selection is linked to both under the cross-sectional design. Types of selection bias commonly encountered by epidemiologists include s*elf-selection bias/loss to follow-up bias*, *nonresponse bias*, *detection bias*, *Berkson's bias*, *healthy control bias*, and *collider-stratification bias*. Two types of selection bias unique to occupational epidemiology are the *healthy worker effect* and the *unhealthy reproducer effect*, which may manifest when occupational groups are compared to the general population. For more detailed explanations and examples, the reader is referred to the sources listed at the end of this chapter.

Confounding bias results from a mixing of the association in which the investigator has interest, with one or more associations in which the investigator does not have interest. This scenario is a consequence of the differential distribution of risk factors among people with and without the study outcome, according to factors other than the outcome under consideration. Confounding is a property inherent to a study population, and factors that confound associations of interest can be identified as those that are common causes of both the risk factor and the outcome of interest. Though not sufficient to cause confounding, a confounding variable necessarily (i) has a causal association with the risk factor of interest in the study population, (ii) has a causal association with the outcome of interest independent of the risk factor of interest, and (iii) is not itself affected by the risk factor or the outcome of interest.

Confounding bias is a nuisance and investigators aspire to eliminate confounding wherever possible. In contrast to information bias and selection bias, statistical adjustment can frequently reduce or eliminate confounding if the relevant factors are identified during study design, and data is accurately captured during study implementation. Approaches to eliminate confounding include exclusion, stratification, standardization, matching, and multivariable analysis, among others. The reader is referred to the sources listed at the end of this chapter for a detailed discussion of these approaches.

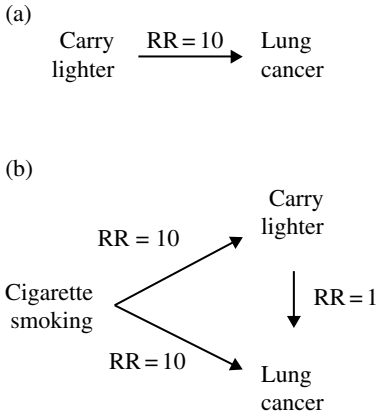


FIGURE 21.4 Example of confounding bias: (a) causal pathway spuriously suggesting that "carry lighter" is associated with a 10-fold increase in "lung cancer" and (b) causal pathway demonstrating confounding of "a" by "cigarette smoking."

When confounding is incompletely addressed, *residual confounding* will continue to bias study results. Age, sex, race and ethnicity, body mass index, socioeconomic status, cigarette smoking, and alcohol consumption frequently merit attention as confounders in environmental and occupational epidemiologic studies.

For the purposes of illustration, Figure 21.4 demonstrates a hypothetical example of cigarette smoking as a confounder of an association between carrying a lighter and lung cancer. In Figure 21.4a, an RR of 10 suggests a 10-fold increase in risk for lung cancer among persons who carry a lighter in their pocket relative to those who do not. However, after adjustment for "cigarette smoking," which is a common cause of "carry lighter" and "lung cancer" in Figure 21.4b, the aforementioned risk disappears as indicated by RR = 1 for the association between "cigarette smoking" and "carry lighter." In this scenario, the association between "cigarette smoking" and "lung cancer" was mixed with the association for "cigarette smoking" and "carry lighter," resulting in confounding.

Effect Modification in Epidemiologic Studies

In contrast to bias, which is a nuisance to be prevented and eliminated if possible, *effect modification* is of interest and cannot be eliminated from a study. Effect modification reflects the biology or other critical component of a risk factor–outcome association. In fact, ignorance of effect modification will counterproductively introduce bias into study results. Effect modification occurs when the association between two or more factors varies by one or more additional factors. An investigator will generally suspect effect modification a priori based on scientific literature, presumed causal pathways between the risk factor and study outcome, or on biology. As a hypothetical example, exposure to an environmental chemical agent might interfere with reproductive success in Asian women and men, yet not in non-Asian women and men; so race is an effect modifier of the agent–reproduction association in this scenario. Sex, age, genetic polymorphisms or markers thereof such as race and ethnicity, and socioeconomic status are frequently encountered as effect modifiers in environmental and occupational epidemiologic studies. The reader is referred to the sources listed at the end of this chapter for further detail.

21.5 EPIDEMIOLOGIC EXPOSURE ASSESSMENT

Exposure entails contact between a person and a risk factor, which is a multifactorial and dynamic process governed by fate–transport processes and toxicokinetics. The goal of an epidemiologic exposure assessment is to delineate groups with various degrees of contact with a risk factor of interest. When the risk factor under consideration in an epidemiologic study is a hazardous or potentially hazardous agent, the exposure analysis represents a critical link for the evaluation of associations with an outcome of interest. Invalid or imprecise exposure assessment introduces information bias and random error into epidemiologic studies. When such error is differential between individuals with a study outcome and those without, internal study validity is compromised. Nondifferential exposure assessment error undermines study power. When exposure in the study sample is not representative of the study population, external validity is compromised. In addition, humans are exposed simultaneously to numerous risk factors in the environment and in the workplace, the effects of which may be *additive*, *synergistic*, or *antagonistic*, and so consideration of exposures to mixtures of risk factors is of increasing concern. The complete mix of all risk factors to which an individual is exposed has been recently defined as the *exposome*. So challenging is the issue of exposure assessment that it is often referred to as the "Achilles heel" of environmental and occupational epidemiology.

As demonstrated by Figure 21.5, there are many factors to consider when assessing exposure, which modify the association between an environmental or occupational risk factor and a study outcome from the point of agent release until an incident outcome. Once released from the source, a risk factor or chemical *agent* in the environment or the workplace is subjected to processes that will affect redistribution in the environment including transport by air, water, soil, climatic factors, and geographic factors, among others, the so-called fate and transport. Agents are deposited in the environment or in the workplace, and contingent on their proximity to persons and their source of deposition, these may come into contact with humans, comprising an exposure. Exposure is a function of (i) the *magnitude* or the

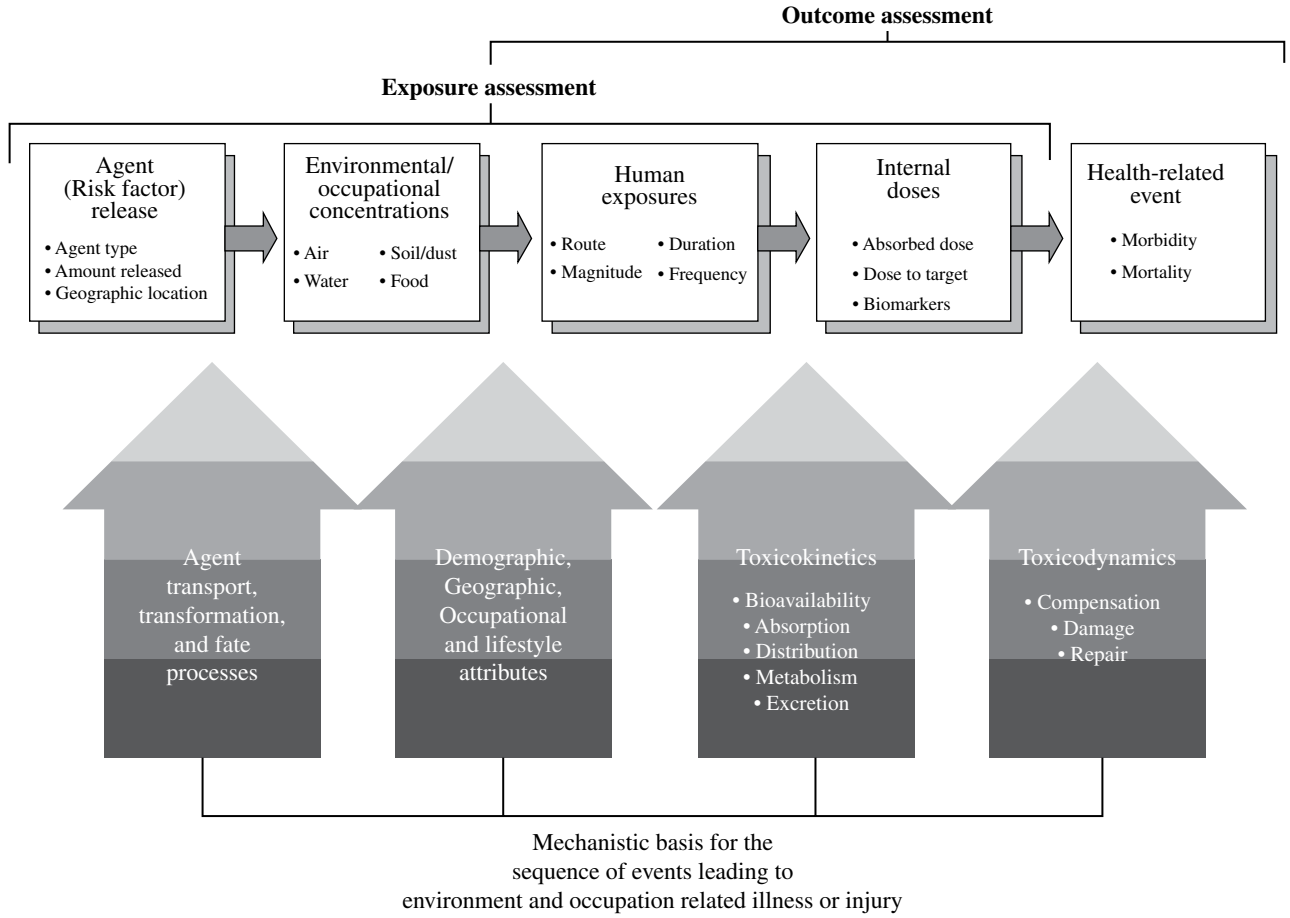


FIGURE 21.5 Mechanistic basis for the sequence of events leading to environmental-related illness or injury. *Source:* Adapted from Sexton et al. (1992). © 1992 Taylor & Francis.

concentration of an agent with which an individual comes into contact (this is often defined in terms of the average exposure over time, the cumulative or total amount of agent over time, or the peak or highest concentration exposure received); (ii) the *frequency* of contact with an agent; (iii) the *duration* of contact with an agent, which may be acute (brief), subchronic (over days or weeks), or chronic (over months or years); and (iv) the *route* of exposure (inhalation, ingestion, dermal absorption, or injection).

Exposure to an agent does not necessarily result in a *dose*, in which an agent penetrates protective epithelial cell layers and gains access to tissues with which it might interact. Dose is often expressed as a concentration, and once inside the human body, *internal dose* is measured. Internal dose is a function not only of the exposure but of the toxicokinetics of an agent including *absorption*, the quantity that enters the body; *distribution*, the volume into which an agent is introduced and sites of accumulation or selective retention; *metabolism*, including phase 1 and phase 2 enzyme detoxification and activation activities; and the rate of *excretion* or elimination from the body. Many factors will modify the toxicokinetics of agents following an internal dose including genetic polymorphisms, nutritional factors, and physiologic or developmental states such as infections, chronic diseases, childhood, and pregnancy. For example, in a large area of Bangladesh, millions of residents are exposed to inorganic arsenic via the consumption of geogenically contaminated drinking water sources. Although well absorbed by the human gastrointestinal tract, there is a wide variability in the effectiveness of arsenic metabolism and consequently in susceptibility to toxic effects associated with genetics, diet, and even pregnancy. The internal dose does not necessarily reflect the quantity of an agent that interacts with a biologic target, such as a phospholipid in a cell membrane, chromatin, or the mitochondrion. The *biologically effective dose* describes the quantity of an agent that interacts with a biologic molecule. Once within the body, agents that are metabolized and excreted at a slower rate than absorbed accumulate in the target organ or tissue. Moreover, a health-related outcome assessed in association with exposure to a risk factor agent is modified by *toxicodynamics*: the adaptation, repair, and compensation that biological systems are likely to employ in an effort to maintain function.

21.6 BIOMARKERS IN EPIDEMIOLOGIC STUDIES

In environmental and occupational epidemiologic studies, it has become essential for epidemiologists to partner with toxicologists to develop new and improved measures of exposure including better internal dose estimates and intermediate markers of biologic effect using *biomarkers*. Toxicologists have been instrumental in development and application of biomarkers in epidemiologic studies. *A Dictionary of Epidemiology* defines a biomarker as a "…cellular, biochemical, or molecular indicator of exposure; of biological, subclinical or clinical effects; or of possible susceptibility …." Environmental and occupational epidemiologists rely increasingly on toxicologists and analytic chemists to develop and evaluate biomarkers for use in the assessment of exposure to hazardous agents, to evaluate outcomes of that exposure, and to identify groups with greater sensitivity or resistance to the deleterious effects of hazardous agents.

Biomarkers are often classified as (i) *biomarkers of exposure*, (ii) *biomarkers of effect*, or (iii) *biomarkers of susceptibility*. However, as demonstrated by Figure 21.6, biomarkers are analogous to "signposts" on a continuum representing the pathophysiology describing the natural history of a study outcome in relation to a risk factor, with the encompassing notion of susceptibility affecting all points along that continuum. These interrelated points on the continuum are also interchangeable in terms of which can serve as a biomarker of exposure to a risk factor and which can serve as a biomarker of an outcome or the event of interest in any given epidemiologic investigation.

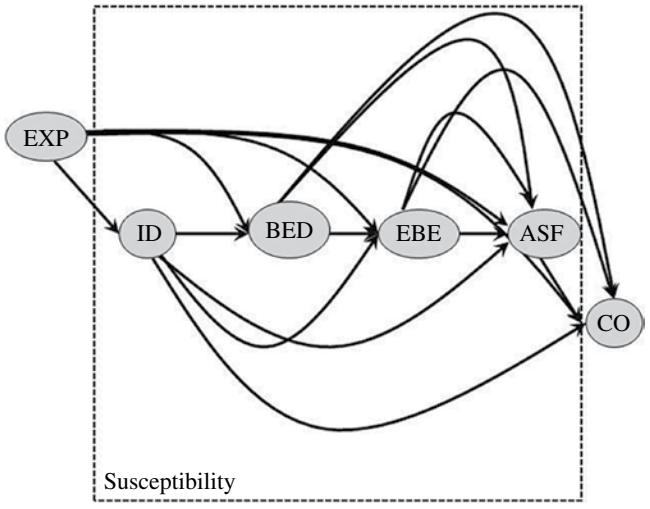


FIGURE 21.6 Biomarkers are interrelated "signposts" along the continuum from risk factor exposure to a clinical study outcome. EXP, exposure; ID, internal dose; BED, biologically effective dose; EBE, early biologic effect; ASF, altered structure and function; CO, clinical outcome. *Source:* Adapted from Schulte et al. (1993). © 1993 American Association for the Study of Liver Diseases.

Biomarker of exposure might include the blood concentrations of an agent of interest such as blood Hg concentrations. Biomarkers of effect might include tissue histology samples from a diseased organ such as skin cancer tissue sections. Biomarkers of susceptibility might include characterization of relevant genetic polymorphisms such as glutathione-*S*-transferase (*GST*) variations or epigenetic modifications of these loci. Their use can help to reduce information bias, in particular for exposure to a risk factor of interest as an investigator may assess the aforementioned internal dose, or biologically effective dose. In addition, biomarkers permit epidemiologists to evaluate subclinical outcomes that precede clinically recognized outcomes, the so-called intermediate effects, including (i) *early biologic effects* and (ii) *altered structure and function.* Whereas such intermediate effects do not comprise a clinical diagnosis or an adverse health outcome per se, these events often precede a study outcome of interest and are likely to occur with greater frequency than a clinical outcome. Early biologic effects are the molecular sequelae of interactions between an agent and its target tissue, such as the F_2-isoprostanes, which are by-products of interactions between reactive oxygen species and fatty acids in cell membranes. Altered structure and function are the morphological and functional changes that occur following interactions between an agent and its target tissue, such as perturbations in thyroid hormone concentrations associated with exposure to environmental chemical agents.

Biomarkers are also useful for assessing the impact of preventive medicine and public health policy efforts, such as the decline in childhood blood Pb concentrations following the U.S. phaseout of tetraethyl Pb from gasoline. *Biomonitoring* projects conduct routine surveillance of exposure to potentially hazardous agents such as metals and organic compounds using biomarker data. Unlike traditional exposure assessments that query study participants concerning behaviors, dietary habits, or work assignments, occasionally augmented by environmental sampling data, collection of biologic specimens and measurement for biomarkers of exposure integrate various sources and routes of exposure such as dietary, respiratory, and dermal transfer. In addition, biomarkers of exposure integrate the toxicokinetics of an agent. Biomonitoring data thus provide invaluable information to environmental epidemiologists and toxicologists in terms of defining *background* exposures to risk factors of interest with which to compare study populations and in identifying potential confounding factors and effect modifiers for epidemiologic investigations.

The use of biomarkers offers many advantages to environmental and occupational epidemiologic studies; however, all biomarkers are not created equal, and a laboratory's ability to measure and report a value for an analyte does not equate with biomarker utility. Biomarkers require verification of accuracy, or *validation*, prior to employment in epidemiologic studies,

and the precision, or *reliability* of a biomarker, is of critical importance in defining its utility. The reader is referred to the sources listed at the end of this chapter for more information.

21.7 HOW EPIDEMIOLOGISTS INFER CAUSALITY

Epidemiology and toxicology work in an integrated fashion to infer *causality* between exposure to a risk factor and a health-related outcome. The detection of an association is not necessarily a reflection of causality; in accordance with the logical fallacy *cum hoc ergo propter hoc*, correlation does not imply causation, though it is indeed supportive. Within a study sample, an association detected between a risk factor and a study outcome may be due to (i) a true underlying causal effect in the study population from which the sample was recruited, (ii) a bias in the approach used to draw the study sample from the study population or in the data collection and analysis strategies, (iii) unmitigated confounding in the study population from which the study sample was recruited, or (iv) random error due to a disproportionate sample having been recruited from the study population. A causal effect cannot be established by the results of any single observational study given the aforementioned limitations inherent to the human condition. Causality in observational studies is inferred under a *weight-of-evidence* paradigm. The results of observational inferential epidemiologic studies are often considered in a hierarchy of sorts with those produced by prospective cohort studies weighted most heavily, followed by retrospective cohort studies, then case–control studies, cross-sectional studies, and finally ecologic studies. Frequently, detection of an association using a cross-sectional or ecologic study design will lead to a more comprehensive and conclusive investigation using a cohort or case–control study design. However, this is not a rule of thumb. As guideposts to assessing evidence for causality, epidemiologists often employ criteria set forth by Sir Austin Bradford Hill during his 1965 President's Address to the Royal Society of Medicine. Sir Bradford Hill provided nine postulates as criteria for the evaluation of causality in observational studies, including:

1. *Strength* of effect, in that the greater the magnitude of an association, the more likely it is that the association is causal in nature
2. *Consistency* of study results, in that an association between a risk factor and an outcome is detected by different investigators, considering various study populations and using varied approaches
3. *Specificity* in the association between a risk factor and an outcome, in that the outcome occurs only following exposure to the risk factor of interest and exposure to the risk factor of interest is not associated with other unrelated outcomes
4. *Temporality* of the association between a risk factor and an outcome, in that exposure chronologically precedes the outcome of interest
5. *Biologic gradient*, or dose–response, such that the greater the extent of exposure to a risk factor, the greater the probability or severity of an outcome of interest
6. *Biologic plausibility*, in which an association between a risk factor and a study outcome can be rationalized or explained in terms of existing toxicologic and biologic knowledge
7. *Coherence*, in that an association does not conflict with what is known in terms of the pathophysiology of the study outcome
8. *Experiment*, in that removal or introduction of a risk factor reduces or increases the frequency of study outcomes
9. *Analogy*, in that comparable associations are reported between similar risk factors and similar study outcomes

With the exception of temporality, no Bradford Hill criterion is requisite to infer causality. Rather, the fulfillment of the criteria strengthens the argument for causality, increasing the weight of evidence for cause and effect between a risk factor and an outcome. For example, in the case of pediatric exposure to ^{131}I and thyroid carcinoma, numerous studies consistently demonstrate an increased risk across various levels of exposure, including background exposure from aboveground nuclear weapons testing in the United States, the Pacific Ocean, and elsewhere, as well as more substantial exposures following the detonation of atomic bombs over Hiroshima and Nagasaki, Japan, and the Chernobyl nuclear reactor meltdown in Ukraine. Despite a moderate increase in risk, increasing pediatric ^{131}I dose corresponds to increasing thyroid cancer risk, and the biologic plausibility for this association is well established. The specificity of the ^{131}I–thyroid cancer association is tenable in that the association appears to be strongest for a particular histologic subtype, and moreover, the association is coherent in terms of radiation and thyroid biology. There is a substantial amount of experimental evidence from animal models and some additional quasiexperimental data contributed by ^{131}I used for medical procedures. Given the aforementioned evidence, the World Health Organization International Agency for Research on Cancer (IARC) has concluded that there exists sufficient evidence in humans and in experimental animals to conclude the existence of a causal association.

To infer a causal association between a risk factor and an outcome, the greatest weight is generally placed on the criteria of temporality, consistency, biologic plausibility, strength, biologic gradient, and experiment. However, exceptions exist in that associations may be present only in

certain groups with enhanced susceptibility to the toxic effect of an risk factor, biological mechanisms may not be appreciated or understood at the time an epidemiologic association is detected, associations between exposures to environmental risk factors and health outcomes may demonstrate only modest increases in risk, and nonlinear or *U-shaped* dose–response curves do not comply with traditional notions of monotonicity. Experimental evidence provides the strongest support for causality. To reemphasize, though temporality is mandatory, the remaining Bradford Hill criteria provide guideposts, rather than requirements, to assessing the critical mass of evidence to infer causal associations using observational data. Recently, the National Research Council of the U.S. National Academies has published an updated risk assessment framework to integrate across multiple lines of evidence including human data, animal data, and mechanistic data.

21.8 SUMMARY

The enormous growth in technology and communications and the exponential increases in sophistication in toxicology and epidemiology over the past two decades increasingly foster an interdisciplinary, collaborative approach to human health studies. The greatest gains in knowledge will result from interdisciplinary team efforts in which various expert team members are not only deeply familiar with their own discipline but also "speak the language" of their colleagues and can communicate intelligibly with expert team members from related disciplines. Nowhere is this truer than in environmental and occupational epidemiologic research. The toxicologist is an integral member of any environmental or occupational epidemiology study team, and participation is necessary at the design, implementation, and data analysis stages. It is our hope that this chapter will foster a greater degree of collaboration between toxicologists and epidemiologists to answer the challenging questions we face and to generate new ones.

SUGGESTED READING

Adami HO, Berry SCL, Breckenridge CB, Smith LL, Swenberg JA, Trichopoulos D, Weiss NS, Pastoor TP. Toxicology and epidemiology: improving the science with a framework for combining toxicological and epidemiological evidence to establish causal inference. Toxicol Sci 2011;122(2):223–234.

Baccarelli A, Bollati V. Epigenetics and environmental chemicals. Curr Opin Pediatr 2009;21(2):243–251.

Bell ML, Samet JM, Dominici F. Time-series studies of particulate matter. Annu Rev Public Health 2004;25:247–280.

Graham S. The sociological approach to epidemiology. Am J Public Health 1974;64(11):1046–1049.

Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. Epidemiology 1999;10(1):37–48.

Hill AB. The environment and disease: association or causation? Proc R Soc Med 1965;58:295–300.

International Agency for Research on Cancer (IARC). Ionizing radiation, part 2: some internally deposited radionuclides. Views and expert opinions of an IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Lyon, 14–21 June 2000. IARC Monogr Eval Carcinog Risks Hum 2001;78(Pt 2):1–559.

Jaffery FN, Misra V, Viswanathan PN. Convergence of clinical toxicology and epidemiology in relation to health effects of chemicals. Environ Toxicol Pharmacol 2002;12(3):169–179.

Kleinbaum DG, Kupper LL, Muller KE, Nizham A. *Applied Regression Analysis and Other Multivariable Methods*. Pacific Grove: Duxbury Press; 1998.

Madia WJ. A call for more science in EPA regulations. Science 1998;282(5386):45.

Morgenstern H, Thomas D. Principles of study design in environmental epidemiology. Environ Health Perspect 1993;101 Suppl 4: 23–38.

National Research Council (U.S.), Committee on Improving Risk Analysis Approaches Used by the U.S. EPA, Board on Environmental Studies and Toxicology and Division on Earth and Life Studies. *Science and Decisions: Advancing Risk Assessment*. Washington (DC): National Academies Press; 2009.

National Research Council (U.S.), Committee on the Institutional Means for Assessment of Risks to Public Health and Commission on Life Sciences. *Risk Assessment in the Federal Government: Managing the Process*. Washington (DC): National Academy Press; 1983.

National Research Council (U.S.), Subcommittee on Arsenic in Drinking Water, Committee on Toxicology, Board on Environmental Studies and Toxicology and Division on Earth and Life Studies. *Arsenic in Drinking Water: 2001 Update*. Washington (DC): National Academy Press; 2001.

Paustenbach DJ, Galbraith D. Biomonitoring and biomarkers: exposure assessment will never be the same. Environ Health Perspect 2006;114(8):1143–1149.

Porta MS, International Epidemiological Association. *A Dictionary of Epidemiology*. 5th ed. Oxford (NY): Oxford University Press; 2008.

Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology*. 3rd ed. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2008.

Sackett DL. Bias in analytic research. J Chronic Dis 1979;32(1–2): 51–68.

Schulte PA, Rothman N, Schottenfeld D. Design considerations in molecular epidemiology. In: Schulte PA, Perera FP, editors. *Molecular Epidemiology: Principles and Practices*, San Diego: Academic Press; 1993. p 159–198.

Sexton K, Selevan SG, Wagener DK, Lybarger JA. Estimating human exposures to environmental pollutants: availability and utility of existing databases. Arch Environ Health 1992;47(6): 398–407.

Szklo M, Nieto FJ. *Epidemiology: Beyond the Basics*. Gaithersburg (MD): Aspen; 2000.

Taioli E, Kinney P, Zhitkovich A, Fulton H, Voitkun V, Cosma G, Frenkel K, Toniolo P, Garte S, Costa M. Application of reliability models to studies of biomarker validation. Environ Health Perspect 1994;102(3):306–309.

Vena JE, Buck GM, Kostyniak P, Mendola P, Fitzgerald E, Sever L, Freudenheim JO, Greizerstein H, Zielezny M, McReynolds J, Olson J. The New York Angler Cohort Study: exposure characterization and reproductive and developmental health. Toxicol Ind Health 1996;12(3–4):327–334.

Woodruff TJ, Burke TA, Zeise L. The need for better public health decisions on chemicals released into our environment. Health Aff (Millwood) 2011;30(5):957–967.

Wu W, Steenland K, Brown D, Wells V, Jones J, Schulte P, Halperin W. Cohort and case–control analyses of workers exposed to vinyl chloride: an update. J Occup Med 1989;31(6):518–523.

22

OCCUPATIONAL AND ENVIRONMENTAL HEALTH

BRENDA S. BUIKEMA, T. RENÉE ANTHONY, AND FREDRIC GERR

The purpose of this chapter is to introduce readers the practice of occupational and environmental health. Specifically, this chapter will define occupational and environmental heath and describe the:

- Historical events that shaped modern occupational and environmental health practices
- Fundamental characteristics of occupational and environmental illness
- Activities and goals of occupational and environmental health specialists
- Multidisciplinary approach to protecting workers' health
- Ethical issues relevant to occupational and environmental health practice

22.1 HISTORY OF OCCUPATIONAL HEALTH

Readers of this book should be familiar with the historical figures whose efforts shaped our modern occupational and environmental health practices as well as the occupational and environmental health disasters that serve as lasting reminders of the importance of preventing occupational and environmental illness and injury.

In 1587, Paracelsus, often considered the father of modern toxicology, wrote the first known monograph on occupational disease. However, the role of occupational risk factors in the development of disease was not widely appreciated until 1700, when the Italian physician, Bernardo Ramazzini, published the first known textbook of occupational medicine, entitled *De Morbis Artificum Diatriba* (*The Diseases of Workers*" in modern translation). Prior to its publication, there were few efforts to link work and illness, despite increasing development of metallurgy and other technologies. Ramazzini recommended that physicians ask their patients about their occupation and the conditions of their work. He also encouraged physicians to visit their patients' work sites to observe directly their working conditions. In particular, Ramazzini wrote, "*Liceat quoque interrogationem hanc adiicere, & quam artem exerceat*" ("I may venture to add one more question: what occupation does [the patient] follow?"). Ramazzini recognized noise, dusts, fumes, extremes of heat and cold, and awkward postures as risk factors for occupational disease and used this knowledge to reduce exposure to these hazards as a means of preventing occupational illness. Ramazzini's work contributed to subsequent efforts of other occupational health specialists and is widely credited with establishing the specialty of occupational medicine.

With the industrial revolution, certain workplace exposures became more common, leading to the occurrence and recognition of new occupational diseases. For example, in 1775, the British surgeon, Percival Potts, noted that occupational exposure to soot among chimney sweeps resulted in an increased risk of scrotal cancer, eventually leading to protective legislation for workers in Europe. Government medical positions were created in the second half of the nineteenth century in Great Britain, establishing occupational medicine as a legitimate part of public health and clinical medicine.

In 1911, the Triangle Shirtwaist Factory fire in New York City resulted in the death of 146 garment workers, which placed workers' health and safety concerns into the limelight both locally and nationally, leading to the reform of fire

Principles of Toxicology: Environmental and Industrial Applications, Third Edition. Edited by Stephen M. Roberts, Robert C. James and Phillip L. Williams.

safety codes and the inception of U.S. workers' compensation programs. This tragedy, resulting from locked fire exits in the factory, also resulted in the creation of the Office of Industrial Hygiene and Sanitation of the U.S. Public Health Service. It was one of the earliest U.S. federal responses to the growing hazards of industrialized work.

Just after World War I, Dr Alice Hamilton, the first woman to be a member of the faculty of the Harvard Medical School, became the leading American figure in the field of occupational and environmental medicine. As a pioneering advocate for occupational and environmental public health, she gave congressional testimony in opposition to the addition of tetraethyl lead to gasoline and correctly predicted widespread environmental lead contamination and an epidemic of lead toxicity as a consequence. In 1925, Dr Hamilton wrote *Industrial Poisons in the United States* and paved the way for growth and maturation of the field.

In 1930, contractor profits were placed ahead of workers' health by allowing dry drilling of silica-containing rock with inadequate ventilation during construct of the 3 mile Hawks Nest Tunnel in West Virginia. The resulting exposure to airborne crystalline silica resulted in the death of hundreds of workers who developed a severe and rapidly progressive form of *silicosis* (a disease of the lung resulting from inhalation of silica dust and characterized by scar tissue and impaired transfer of oxygen from air to blood). Hundreds more developed permanent lung disease. In terms of lives lost, this incident remains the single largest industrial disaster in U.S. history.

In December of 1952, a 5-day temperature inversion in London, United Kingdom, resulted in a dense ground-level fog that trapped soot, tar, and sulfur dioxide from household and industrial coal combustion. London residents developed chest pain, lung inflammation, and increased asthma incidents, with 4000 fatalities initially attributed to the event. While many who died of respiratory illness during the inversion were elderly or had preexisting lung conditions, increased fatality rates for both middle-aged and infant residents were observed during and immediately following the smog event. Today, estimates of 12,000 deaths are attributed to the event. In the wake of this tragedy, the British government began efforts to reduce household coal emissions in the ensuing 1956 Clean Air Act.

Such industrial disasters have affected not just previous generations but are still occurring in the current era. For example, in 1984, the unplanned release of methyl isocyanate gas from a Union Carbide plant into a densely populated neighborhood of Bhopal, India, resulted in the immediate death of approximately 2000 people and the onset of permanent lung disease among more than ten thousand residents. Neither the community nor local hospitals knew what was in the gas or were aware of its toxic effects. As a result, regulators in the United States developed a hazard communication standard to ensure that the hazards of occupational chemicals were clearly communicated to workers and, in conjunction with emergency response rules, to the neighboring communities and emergency responders. In 1986, an explosion at the Chernobyl nuclear power plant in Ukraine resulted in the immediate deaths of 28 workers and emergency responders, who again did not know the hazards of the radiation exposures they were facing. In 2005, an explosion in a British Petroleum refinery in Texas City killed 15 workers and injured 170 others because of violations of the process safety management standard, and in 2010, 29 coal miners died in a mine explosion at the Upper Big Branch Mine in West Virginia because inadequate ventilation allowed explosive dusts and gases to build in the mine.

While these cases illustrate the diversity of large-scale tragedies associated with industrial exposures, it is important to remember that many more workers are chronically exposed to hazardous chemicals that may be associated with health effects that may take years to develop.

Over the ensuing decades, multiple occupational and environmental health outbreaks have occurred in the United States and throughout the world. In the United States, the Occupational Safety and Health Act of 1970, which created both the Occupational Safety and Health Administration (OSHA) and the National Institute for Occupational Safety and Health (NIOSH), allowed for greater oversight and regulation of occupational exposures and furthered the field of occupational and environmental medicine in the United States. As a result of the contributions of both early and recent occupational health researchers and practitioners, the health and safety of workers has improved dramatically, especially in developed nations.

22.2 DEFINITION AND SCOPE

Industrial and commercial activities create a wide range of human health hazards in both the workplace and the general environment. In the workplace, exposure to machinery, toxic chemicals, physical agents (e.g., noise, heat, and radiation), and biological hazards result in illness and injury. Releases of toxic substances into ambient air and water have adverse environmental and human health effects. When manifesting as a well-defined human disorder, the adverse effect of an exposure to an occupational and environmental hazard is typically categorized as an *injury* or *illness*. An occupational injury is an adverse health event of nearly instantaneous onset resulting from a single, short-duration occupational exposure.[1] Examples include a burn experienced by a worker who has contact with steam escaping from a ruptured pipe

[1] Although nonoccupational injury (e.g., those resulting from automobile crashes) is an important public health problem in both developed and developing societies, it is typically studied and regulated separately from occupational injury.

TABLE 22.1 Examples of Organ System, Occupation or Setting, Specific Exposure, and Health Effects for Selected Occupational Diseases

Organ System	Occupation or Setting	Exposure	Disease
Respiratory	Shipyard, insulation installer/ remover	Asbestos	Pulmonary fibrosis, lung cancer
	Foundry, sand casting	Silica	Pulmonary fibrosis, lung cancer
	Agriculture, farmer	Organic dusts	Hypersensitivity pneumonitis
	Painting and coatings	Toluene diisocyanate	Asthma
	Coal mining, coal dust	Coal dust	Coal workers' pneumoconiosis
Neurological	Automobile repair, mechanic	Organic solvents	Cognitive impairment
	Storage battery manufacturing	Lead	Kidney disease, neurological disease, anemia
	Construction, agriculture, others	Noise	Hearing loss
Dermatological	Healthcare workers, agricultural workers, others	Soaps, solvents, sensitizing agents, physical agents	Contact dermatitis, skin cancer
Hematological	Petroleum refining	Benzene	Anemia, leukemia
Kidney and bladder	Plating, battery manufacturing	Cadmium	Renal impairment
	Tanning and dying	Aniline dyes	Bladder cancer
Cardiovascular	Rayon manufacturing	Carbon disulfide	Atherosclerosis

and a broken leg experienced by a worker who falls from a roof. Of more interest to the study of toxicology, injuries from chemical exposures can also occur. Chemical injuries associated with short-term exposures typically require inhalation or dermal exposures to high concentrations of acutely hazardous chemicals. Examples include exposures to (i) airborne hydrogen sulfide, where a single exposure to concentrations above 700 mg/m^3 can cause rapid respiratory failure; (ii) dermal contact with phenol, which is rapidly absorbed through the skin and can result in irregular breathing, muscle weakness, and respiratory arrest; and (iii) skin contact with concentrated hydrofluoric acid, which causes both initial redness and pain and can result in hypocalcemia (low blood calcium) and death following contact with as little as 2.5% of the body's skin.

An occupational or environmental illness is an adverse health condition that occurs over some period of time after exposure to a hazardous agent. Examples of occupational and environmental illness include hearing loss among workers who experience prolonged exposure to noise, asthma (an obstructive lung disease) among workers with occupational exposure to the organic chemical toluene diisocyanate, and impaired cognitive function experienced by a child who ingests lead-based paint found in older housing. None of these illnesses occur immediately following the initial exposure.

Occupational illness can result from exposure to chemicals (e.g., pesticides from agricultural applications, solvents from printing operations, and metals from welding), minerals (e.g., asbestos from mining, silica from abrasive blasting, coal dust from mining), physical agents (e.g., vibration from construction equipment, radiation from radon intrusion), biological agents (e.g., hepatitis B from tattooing, tuberculosis from hospital patients, and human immunodeficiency virus from needlestick injuries), or psychosocial stressors (e.g., time and productivity demands from automated production lines). Some examples of the large variety of work-related diseases encountered by occupational medicine providers are provided in Table 22.1.

The goals of occupational and environmental health are to prevent, identify, and mitigate the adverse effects of occupational and environmental hazards among workers and exposed members of the public.

22.3 DATA SOURCES AND THE BURDEN OF OCCUPATIONAL AND ENVIRONMENTAL ILLNESS AND INJURY

Information on occupational injury and illness is available from several sources, each providing a partial characterization of the full impact of these conditions. The Bureau of Labor Statistics (BLS) of the U.S. Department of Labor is the primary federal entity responsible for collection of work-related illnesses and injuries statistics in the United States (www.bls.gov/). BLS occupational injury and illness information is obtained by compilation of injury and illness reports submitted by eligible employers as required by applicable U.S. OSHA law. There are limitations to these data, however. First, many employers are exempt from federal reporting requirements (e.g., farms with fewer than 11 employees and self-employed persons). Further, underreporting of injuries and illnesses by employers has been identified across many industrialized nations. Underreporting of occupational disease is common due to the chronicity and latency of many occupational diseases (discussed in the following text). In addition to the BLS, occupational illness and injury information can also be obtained from state and

federal workers' compensation reports; however, substantial variability in compensation programs across states limits the utility of this information. As with BLS data, long latent interval diseases from occupational exposures may be neither recognized nor captured by worker's compensation systems.

The data that are collected can be analyzed by industry, exposure type, and reported adverse health effect. Ideally, this information is used for prevention efforts. For example, the OSHA recently issued an alert to workers in the salon industry to notify them of excessive levels of formaldehyde in professional hair care products (http://www.osha.gov/SLTC/formaldehyde/hazard_alert.html).

Despite limitations in its completeness, valuable information about the pattern and distribution of work-related injuries and illnesses can be obtained from BLS data. According to 2009 BLS data, 1,238,490 people experienced an injury or illness at work in private industry and state and local government agencies that required one or more days away from work. In private industry in 2009, 14,350 injuries and illnesses requiring time off from work were attributed to exposures to chemicals and chemical products. The breadth of illness and the extent of workers with these illnesses provide unique challenges to all occupational and environmental health specialists.

22.4 CHARACTERISTICS OF OCCUPATIONAL ILLNESS

Illness resulting from occupational and environmental exposures is often overlooked and underreported. This is unfortunate, since neither prevention nor effective treatment is possible unless the link between the exposure and the illness is recognized. A discussion of special characteristics of occupational and environmental illness is provided in the following text that, when understood, will assist in better recognition and prevention of these disorders.

The Biological, Clinical, and Pathological Characteristics of Occupational Illness Are Often Indistinguishable from Those of Illness of Nonoccupational Origin

Simply stated, the clinical and biological characteristics of a disease that results from occupational or environmental exposures are no different than those of the same illness resulting from other causes. Examples are numerous—lung cancer resulting from occupational exposure to asbestos is identical to lung cancer resulting from exposure to tobacco smoke or radon gas. Anemia from occupational exposure to lead may be identical to that resulting from iron deficiency or other diseases. Hearing loss caused by occupational exposure to noise is indistinguishable from hearing loss resulting from nonoccupational exposure to noise. Perhaps more than any other single fact, the similarity of occupational illness to nonoccupational illness makes it especially important to maintain a high level of vigilance (in medical terms, a high "index of suspicion") for occupational causes of common illness.

The Cause of Many Occupational Illnesses Is Multifactorial

Occupational and nonoccupational factors can act together in the causation of disease. For example, regular smokers of cigarettes have about a 10- to 15-fold increase in lung cancer risk. Persons with a history of occupational exposure to asbestos have about a fivefold increase in lung cancer risk. However, among persons exposed to both cigarettes smoke and asbestos, the risk of lung cancer is 50 or more times that of persons who have no exposure to either hazard. In this example, we see that the risks are not simply additive, but actually multiply by each other. Such synergy of the effect of multiple exposures has been well characterized for only a few combinations of exposure, but likely occurs often.

The Adverse Effects of Occupational Exposure May Begin after a Predictable Interval from Onset of Exposure

The adverse health effects of specific exposures may not be observed for many years after cessation of the exposure. The time period separating the onset of exposure to the clinical manifestation of illness is called a "latency period." For example, some persons exposed to airborne silica dust may not manifest scarring of the lung (i.e., silicosis) until more than a decade has passed (although it can occur both earlier and later than 10 years). Latency periods vary greatly across toxicants and illnesses. For example, organophosphate pesticide poisoning may occur within minutes or hours following exposure, whereas cancers resulting from occupational and environmental exposures (e.g., lung cancer from occupational exposure diesel exhaust) occur after a latency period of years to decades.

The Dose of a Toxicant Is an Important Predictor of Health Effects

As emphasized elsewhere in this book, the dose of a toxicant absorbed by an individual is a critical determinant of the resulting adverse health effects. For exposure–effect associations that are *stochastic* in nature (i.e., probabilistic), the dose is a primary determinant of the risk of disease. Stochastic dose–response relationships are observed for carcinogenic agents—as the dose of the carcinogen increases, the probability of developing cancer increases (however, the severity of the cancer itself is not a function of the dose).

Alternatively, for *deterministic* exposure–effect associations, the dose affects the severity of the disease. For example, the severity of the acute adverse effects of organophosphate pesticide exposure is highly dependent on the dose received. Virtually, all exposed persons, at sufficient doses, will exhibit well-characterized signs of toxicity.

People Vary Substantially in Their Responses to Occupational and Environmental Exposures

Considerable variability in both risk and severity of disease is observed among exposed persons even after dose has been considered. Such variability may be due to other environmental exposures in both work and nonwork settings (as noted earlier), individual characteristics (e.g., physical conditioning, nonoccupational health behaviors, and overall health status), and differences in genetic makeup across individuals. For example, workers who are exposed to a hepatotoxic agent such as carbon tetrachloride are at higher risk of liver toxicity if they are also regular consumers of ethyl alcohol, which is also known to be toxic to the liver.

22.5 THREE PREVENTION GOALS

In order to maintain the health and productivity of working people, occupational health providers are committed to three major prevention goals: (i) prevent hazardous exposures from occurring to minimize or eliminate the risk of occupational and environmental disease, (ii) identify early evidence of harm (i.e., preclinical disease) in order to prevent additional exposure and further harm, and (iii) diagnose and treat diseased individuals to prevent further health deterioration. Specific activities of occupational and environmental health professionals are described in the following text for each of these goals.

The first prevention goal is called *primary prevention*. Primary prevention is accomplished by lowering the risk of an occupational or environmental disease among healthy (i.e., disease-free) persons. To do so, occupational health professionals must recognize hazardous exposures and take steps to control them before resultant health consequences occur. Primary prevention in the workplace is often implemented by occupational health professionals specializing in *industrial hygiene*. Industrial hygienists are trained to anticipate, recognize, evaluate, and control exposures in the workplace to prevent sickness, impaired health and well-being, or significant discomfort among workers. For example, after recognizing hazardous airborne mercury concentrations in a work environment, industrial hygienists will identify and characterize the source of the vapor and work with engineers to find a safer substitute, change the facility ventilation, or otherwise contain the substance.

Industrial hygienists have prioritized exposure reduction approaches into a hierarchy of controls, described in Table 22.2. The most effective control measures focus on elimination or reduction of the hazard with methods that require no change to workers' behavior. Administrative controls, such as implementing work/rest cycles on hot days, require supervision reinforcement and workers' participation to minimize health risks. The final level of control, that is, use of personal protective equipment (e.g., respirator masks), relies on the worker to properly use and maintain protective equipment. While these are often effective to reduce exposures, they place the burden of protection on the worker and are usually considered the last line of defense for primary prevention. Primary prevention in occupational and environmental health may also be implemented by substituting less hazardous chemicals for ones in current use.

TABLE 22.2 Hierarchy of Control

Control Option	Description
Primary control: engineering	
Elimination	Eliminate the hazardous chemical use or overall process
Substitution	Replace the hazardous chemical with a less toxic chemical
Isolation	Enclose the process/locate process in unoccupied areas
Automation	Use automated equipment to transfer chemicals; use robotic equipment to perform worker activities
Ventilation	Increase general air movement in occupied spaces; provide localized ventilation where hazardous chemicals are generated
Secondary control: administrative	
Job rotation/supervision	Minimize the duration of exposure by rotating workers throughout the shift; supervision is critical to minimize chronic exposures
Tertiary control: personal protective equipment	
Respiratory protection	Provide certified respirators approved for the hazard of concern (must comply with regulatory standards)
Chemical protective clothing	Provide chemical-resistant clothing to prevent contact with hazardous chemicals (must select materials based on resistance to chemical(s) in use)

For example, over the past several decades, safer water-based latex paints have been used as a substitute for more hazardous solvent-based paints in residential and commercial building construction. This trend has substantially decreased exposure to organic solvents among construction painters. Substitution does not always eliminate health hazards of a particular industrial sector, however, as the new chemical may introduce unexpected hazards. For example, the commercial dry cleaning industry has focused on replacing perchloroethylene, a chemical hazardous to the nervous system and suspected of being cancer-causing agent, with 1-bromopropane (1-BP). However, workers handling the new chemical also exhibited neurological effects, including light-headedness and loss of sensation in the arms and legs (peripheral neuropathy). NIOSH investigations identified that exposures to 1-BP exceeded recommended exposure limits and recommended additional ventilation to reduce worker exposure. While the hazards of the substituted chemical, 1-BP, were well documented, sometimes, chemicals with ingredients of unknown toxicity are substituted for chemicals with known toxicity. In these cases, industrial hygienists must work with toxicologists to perform hazard assessments with appropriate analogous chemicals to anticipate hazards of substitution chemicals.

Sometimes, to minimize exposure, substitution or other exposure reduction controls are not feasible, and administrative changes are recommended, instead. For example, workers who are exposed to repetitive movements of their wrists and hands for 8 h per day may be instructed to rotate to other jobs to reduce the workers' risks of musculoskeletal conditions. It is important to recognize that such efforts are generally not as effective as reducing or eliminating exposure at the source. Finally, if exposure is not well controlled by engineering or administrative controls, personal protective equipment (e.g., respirators, gloves, earplugs) is often used. It is important to remember that personal protective equipment is rarely as effective as substitution of less toxic products or engineering approaches.

The next goal of occupational health practice is *secondary prevention*. Secondary prevention is implemented by screening for subtle disease states or for deviation from health early in its evolution, prior to overt clinical manifestation of illness (i.e., "early detection"). Early detection allows for the control of exposure or treatment of the health condition before more serious illness occurs. An example of secondary prevention is the monitoring of nickel–cadmium battery-manufacturing workers for evidence of the adverse effects of cadmium exposure on the kidney. When cadmium exposures are confirmed or suspected, biological measures of (i) cadmium absorption (i.e., cadmium concentration in the blood and urine) and (ii) early adverse effects of cadmium on the kidney (i.e., measurement of the protein beta-2 microglobulin in the urine) are often performed. Workers experiencing subtle adverse renal effects of cadmium exposure have elevated concentrations of beta-2 microglobulin in their urine despite the fact that no disease has yet occurred. These elevated concentrations indicate the need for removal of the worker from cadmium exposure to prevent the development of more advanced renal disease [29 CFR 1910.1027(L)(3)]. Another example of secondary prevention is provided in the OSHA lead standard [29 CFR 1910.1025]. The standard requires a medical surveillance program for all employees who are, or may be, exposed to lead at levels higher than the OSHA-designated action level for more than 30 days per year. As part of the program, medical examinations, blood lead levels, and laboratory tests of blood and urine are monitored. If the blood lead levels are elevated, then the medical surveillance program guides the healthcare provider on the appropriate action to be taken with the worker (i.e., removal from the environment until the worker's blood lead level decreases).

The third prevention goal of occupational health practice is *tertiary prevention*. Tertiary prevention is the diagnosis, treatment, and rehabilitation of a clinically overt occupational or environmental disease process with the goal of limiting the severity of the condition or its long-term consequences. An example of this would be recognition of occupational asthma and the subsequent treatment of the worker to prevent excessive inflammation and scarring of the air passages in the lungs.

22.6 ACTIVITIES OF OCCUPATIONAL AND ENVIRONMENTAL HEALTH PROVIDERS

Occupational and environmental health is a multidisciplinary field and is served by a wide range of health professionals. Each provider has specific educational requirements and professional certification, and typically, no one provider is sufficiently skilled to implement all of the components of a successful occupational and environmental health program. The major professional disciplines in occupational health are described in the following text, along with specific roles in illness prevention.

Occupational Medicine Physician

Occupational medicine physicians provide medical services to workers by providing routine examinations, performing medical surveillance testing, and evaluating and treating occupational illness and injury. These physicians are certified by the American Board of Preventive Medicine, and many have completed occupational medicine residencies, obtaining master's degrees in public health. Occupational medicine physicians work in multiple settings. In the past, many worked in large industrial facilities and were employed by a single company. More recently, they are likely to be a partner in an independent occupational medical practice that

provides occupational medicine services under contract to many companies.

Occupational medicine physicians provide both routine scheduled clinical services (i.e., preplacement examinations given before the start of work and medical surveillance examinations given periodically to workers with potentially hazardous exposures) and care for unplanned or unexpected injury and illness. Preplacement examinations are often given to workers who are going to work in high-noise areas (audiometric testing), who are required to wear respiratory protection (pulmonary function testing), and who perform emergency response activities. Employers in the United States are also authorized to engage in preplacement examination of other health parameters, such as physical strength. However, employers are limited in their use of this information by laws that protect against discrimination of disabled persons. Overall, the purposes of the preplacement examination include (i) assessment of whether the employee is physically able to perform all essential functions of the job, (ii) evaluation of the potential for an employee's health to put others at risk of harm, and (iii) assessment of the potential for an employee's health to put himself/herself at risk of harm as a result of specific exposures and conditions expected on the job.

Medical surveillance examinations given periodically to workers with potentially hazardous exposures are a critical component of *secondary prevention*. The OSHA specifies minimum medical surveillance examination requirement in numerous chemical-specific standards (e.g., 29CFR1910.1001–1096), hazardous waste operator standard (29CFR 1910.120), and respiratory protection standards (e.g., 29CFR1910.134). For example, on an annual basis, employers in the United States must provide workers routinely exposed to asbestos with an opportunity to be evaluated by a medical provider. As part of this evaluation, the employee will be asked questions about his or her breathing as well as be offered a physical examination. The worker will be evaluated with pulmonary function tests (using *spirometry*) to provide information about lung function and a chest X-ray to reveal early changes associated with asbestos exposure. These tests are performed on a regular basis so that those workers who are exhibiting early signs of disease can be removed from exposure. There are many substances and hazards for which the OSHA has mandated medical surveillance standards. Examples of these standards include monitoring for health effects of arsenic, asbestos, benzene, cadmium, lead, vinyl chloride, and noise exposures. Each agent has different requirements for physical examination and laboratory testing, and the occupational health provider should be knowledgeable of the specific details of them.

Occupational medicine physicians have a central role in *tertiary prevention* by diagnosing and treating occupational injuries and illnesses. Examples include medical providers suturing a laceration of the skin caused by contact with a cutting tool or diagnosing and treating asthma in a patient who complains of intermittent cough and shortness of breath while at work. After making a diagnosis and offering treatment, the provider typically also recommends work restrictions (i.e., limitations to specific occupational activities or exposures) to limit future harm.

Occupational medicine physicians may also be called upon to offer an opinion whether an individual worker's disease was caused by a specific occupational exposure. Often, these cases are linked to monetary settlements, either through the workers' compensation system or during a liability lawsuit. Occupational and environmental medicine providers who participate in these evaluations must have extensive knowledge of the patient's work and exposure history, medical history, and current health status. In addition, they must have knowledge of the medical literature supporting or refuting a linkage between the exposure and disease outcome. These cases are often complex and intellectually challenging.

Finally, occupational physicians perform the important role of providing education to workers. One-on-one education about health hazards, exposure risks, and the risk of illness from work-related and nonwork-related exposures is essential to each interaction with a worker.

Occupational Health Nursing

Occupational health nurses (OHNs) provide health and safety programs and services to workers, worker populations, and community groups. The practice focuses on prevention of occupational illness and injury and protection from work-related and environmental hazards. In addition, OHNs often engage in wellness and health promotion activities. They do so by engaging in the three levels of prevention described earlier. The services that OHNs provide include (i) initial clinical treatment of occupational injury or illness; (ii) case management, that is, the process of coordinating and providing oversight of all aspects of treatment and rehabilitation of individuals with occupational illness or injury; (iii) health promotion, which includes counseling or referral of workers to manage behavioral (e.g., smoking) or modifiable conditions (e.g., high blood pressure) that increase risk for adverse health outcomes; (iv) hazard detection and mitigation, often in collaboration with an industrial hygienist or a safety specialist; and (v) legal and regulatory compliance with multiple laws and regulations affecting employment and the work environment. Nurses can become a certified OHN through the American Board for Occupational Health Nurses. Currently, the United States has 12,000 OHNs.

Industrial Hygiene

As discussed previously, industrial hygienists are integral to the *primary prevention* of exposure-related disease in the workplace. Industrial hygienists are familiar with industrial

processes, tools, and equipment and have the skills necessary to evaluate, quantify, and mitigate hazardous exposure. Industrial hygienists apply knowledge in basic sciences (e.g., chemistry, physics, biology, and toxicology) to conditions in workplaces to monitor, assess, and control exposures that may cause worker injuries and illnesses.

Industrial hygienists specialize in techniques to monitor worker exposures to chemical and physical agent hazards (e.g., noise, heat, and radiation) and to interpret these results to determine the risk of illness. Risk evaluations are conducted to determine whether exposures are sufficiently below regulatory limits and consensus standards. To assess compliance with state or federal OSHA standards, exposure measurements are compared to OSHA permissible exposure limits (PELs), found in 29CFR1910.1000 (for general industry, commonly referred to as the "Z-tables") as well as chemical-specific standards (29CFR1910.1001 through 1096). The PELs in the Z-tables were established at the inception of the OSHA in 1972, whereas chemical-specific standards have been introduced and updated over time. Because they are updated more frequently, industrial hygienists more commonly assess exposure risks provided in consensus guidelines, which are not legally enforceable by the OSHA but are based on more recent toxicological and epidemiological data. These include the NIOSH recommended exposure levels (RELs) and the American Conference of Governmental Industrial Hygienists (ACGIH) threshold limit values (TLVs™). The ACGIH TLVs are routinely reviewed and establish maximum concentrations that "nearly all workers" may be exposed to repeatedly, throughout a working lifetime, without adverse effect. The NIOSH, the agency responsible for recommending standards to OSHA and the Mine Safety and Health Administration (MSHA), develops RELs based on determinations of health risks from human or animal data and on assessments of levels that can be feasibly achievable. Previous to 1995, feasibility was not included in NIOSH RELs, but current updates must consider whether engineering controls can achieve these levels and whether analytical techniques are available to assess exposures below an established REL.

For example, a work site using acetone has to legally comply with the OSHA PEL of 1000 ppm (from the Z-tables). However, this value was based on the 1968 ACGIH TLV, which has been updated over the decades and now is set to 500 ppm to prevent adverse neurological effects. But the NIOSH has also established an 8 h REL for acetone at 250 ppm. Simply complying with an OSHA PEL from the Z-tables will *not* protect the health of the workforce, so an internal occupational exposure limit that sets the maximum exposure to all possible exposures at a workplace is often established. A good industrial hygienist evaluates the documentation that was used to generate the associated exposure limit to determine which exposure limit is most appropriate to protect workers' health. In many cases, there may be exposures or health effects associated with compounds for which there is no published exposure limit, and the hygienist must develop internal recommendations based on toxicological principles.

In addition to evaluating workplace exposures to hazardous chemicals and physical agents, industrial hygienists select, fit, and evaluate the performance of chemical protective clothing and respiratory protection. They are also skilled in evaluating the performance of and recommend changes to engineering control systems to remove contaminants where they are generated (local exhaust ventilation) or improve mixing and airflow in large work areas do dilute less hazardous contaminants in the workplace (general ventilation). Hygienists also develop and deliver chemical-specific hazard communication training, assist with preplanning emergency response teams, and work closely with occupational medicine physicians and OHNs.

After specialized education and work experience, these occupational health professionals can become a certified industrial hygienist (CIH) through the American Board of Industrial Hygiene (ABIH) or, in Canada, registered occupational hygienist (ROH) through the Canadian Registration Board of Occupational Hygienists. Currently, there are nearly 8000 CIHs worldwide.

Ergonomists

Ergonomics is the study of the relationship between the physical demands of work and the ability of the worker to meet those demands over time. Ergonomists examine the interaction of the physical workplace with the human body. Specifically, ergonomists evaluate and quantify the physical demands of the job, based on human performance needs for specific work tasks and the tools and equipment in use for completing that task. They work to design equipment, products, facilities, devices, tasks, and systems to minimize injuries and illnesses of the musculoskeletal system. The goal of ergonomics is often stated as "fitting the job to the worker rather than the fitting the worker to the job." Ergonomists are often asked to guide employers who want to control physical hazards for low back pain and upper limb disorders (such as rotator cuff tendonitis and carpal tunnel syndrome). To reduce the risk of low back pain, ergonomists design work with less frequent and forceful lifting, bending, twisting, and whole-body vibration (e.g., as experienced by truck drivers) experienced by workers. To reduce risk for upper extremity disorders, ergonomists design work processes with lower forces and frequencies of repeated upper limb motions and by arranging workstations and tools so that joints are not used at the extremes of their ranges of motion. Professional certification for ergonomists (CPE) is awarded through the Board of Certification in Professional Ergonomics in the United States and the Canadian College for the Certification of Professional Ergonomists (CCCPE) in Canada. There are currently 1500 certified ergonomists.

Other Specialists

Most manufacturing locations have a designated safety manager who administers many aspects of workers' health and safety. While they typically rely on industrial hygienists to focus on assessing health risks from exposures, such safety specialists are integral to educating workers on many health hazards and organizing ongoing health and safety initiatives with workers. Approximately 11,000 certified safety specialists (through the Board of Certified Safety Professionals) are employed in the United States.

Other specialists involved in occupational health include toxicologists, engineers (industrial and environmental), epidemiologists, and health physicists (radiation hazard specialists). These specialists are routinely consulted to provide technical expertise to previously described occupational health practitioners and are integral to successfully understanding health risks to workers and the general public.

Many occupational health technicians are also integrated into successful occupational health programs. The NIOSH provides approval for spirometry training courses, recommended for pulmonary function testers who are not licensed health professionals. The Council for Accreditation in Occupational Hearing Conservation (CAOHC) provides certification for audiology technicians to perform standard hearing testing (under the supervision of a professional audiologist supervisor). These technicians, among others, are useful to performing frequent, standard tests that are often required by occupational health regulations.

22.7 THE MULTIDISCIPLINARY APPROACH TO WORKERS' HEALTH

As illustrated in the previous section, the skills required to successfully protect workers' health require a team of experts with distinct skills who work together to maximize the health of the working people. It is rare that a full complement of providers is hosted at a given work site. The resources available at a given work site vary, typically depending on the size of a given work organization and the type of hazards that are present (Table 22.3). At a small facility, much of the day-to-day occupational health activities may be performed by a safety manger (with or without a CSP) or, in some cases, a human resources manager. At larger facilities, the safety manager is usually certified and the plant may employ an OHN: this team is typically responsible for managing injury/illness cases as well as proactive aspects of health and safety programs. For large national or multinational companies, there may be a corporate safety and health department and a corporate medical director: these groups establish internal company policies (such as which exposure limits to adopt and company-specific medical evaluation tests). Members of the corporate health staff often travel to multiple manufacturing sites to provide training, measurements, and interpretation of cases to assist local occupational health staff.

In recent times, a majority of manufacturing locations contract many occupational health services, including occupational health physician services, audiology, ergonomics, and industrial hygiene monitoring. For successful occupational health programs, it is critical for contracted services to be aware of specific hazards their clients' workplaces. For example, a physician needs to have a specific description of work duties, required personal protective equipment, and environmental hazards to determine whether an injured/ill worker can return to a particular job. A contracted audiologist must include noise exposure data on their hearing test reports to assess work-related hearing loss. An ergonomist must understand where musculoskeletal injuries are most common to determine where to focus an investigation and then must understand how workers interact with tools, equipment, and product to assess risk and recommend interventions. In short, occupational health professionals,

TABLE 22.3 Examples of Occupational Health Staff Structure By Facility Size and Risk

Size	Nature of OEH[a] Risks	On-Site Personnel	Corporate Assistance	Contracted Services
Small	Minimal chemical/noise hazards	Safety manager/human resources	None	IH monitoring, medical, nursing
	Chemical/noise	Safety professional, human resources	Access to corporate IH, safety	IH monitoring, medical, nursing, audiology
Medium	Minimal chemical/noise hazards	Safety professional	Access to corporate IH, safety	IH monitoring, medical, nursing
	Chemical/noise	Safety professional, nurse	Corporate IH, safety, medical	IH monitoring, medical, nursing, audiology
Large	Minimal chemical/noise hazards	Safety professional	Access to corporate IH, safety	IH monitoring, medical, nursing, audiology
	Chemical/Noise	Industrial hygienist Safety professional Nurse	Access to corporate IH, safety, medical	Medical

[a]OEH, occupational and environmental health.

whether contracted or employed at the facility, must work as a team to fully characterize health hazards and to prevent worker illnesses and injuries that may result from them.

Even with the diversity in occupational and environmental professional staffing at manufacturing locations, a common set of tasks are needed to protect worker heath. Specifically, the staff must identify and assess physical, chemical, and ergonomic hazards; quantify exposures and estimate risk based on exposure quantification; evaluate and recommend appropriate controls; and evaluate workers' health status. Figure 22.1 provides a generic template for understanding a general process for completing each of these tasks, illustrating the interrelatedness of the health paradigm. Tables 22.4 and 22.5 illustrate how each of the occupational health professionals may contribute to completing these tasks for two specific hazards that may be identified at a facility, i.e., chemical and lifting exposures, respectively. These tables highlight the interdisciplinary nature of occupational health protective services and how the resulting analysis from each professional is needed to complete a comprehensive risk assessment to ensure adequate health protection.

22.8 LEGAL AND REGULATORY ISSUES PERTINENT TO OCCUPATIONAL AND ENVIRONMENTAL HEALTH

The two U.S. federal agencies of the United States primarily responsible for the health and safety of workers are the OSHA, a unit of the of the U.S. Department of Labor, and the NIOSH, a component of the Centers for Disease Control and Prevention (of the Department of Health and Human Services). Both organizations were created by passage of the Occupational Safety and Health Act of 1970. The OSHA's mission is to "ensure safe and healthful working conditions for working men and women by setting and enforcing standards and by providing training, outreach, education and assistance." The OSHA is empowered to establish legally binding upper limits for exposure to occupational hazards (PELs) and to enforce such standards through on-site workplace inspections. If violations are noted, OSHA is empowered to fine the employer an amount based on severity of the violation. The NIOSH assists the OSHA by (i) providing the agency with scientific research on the human health effects of workplace exposures and (ii) making

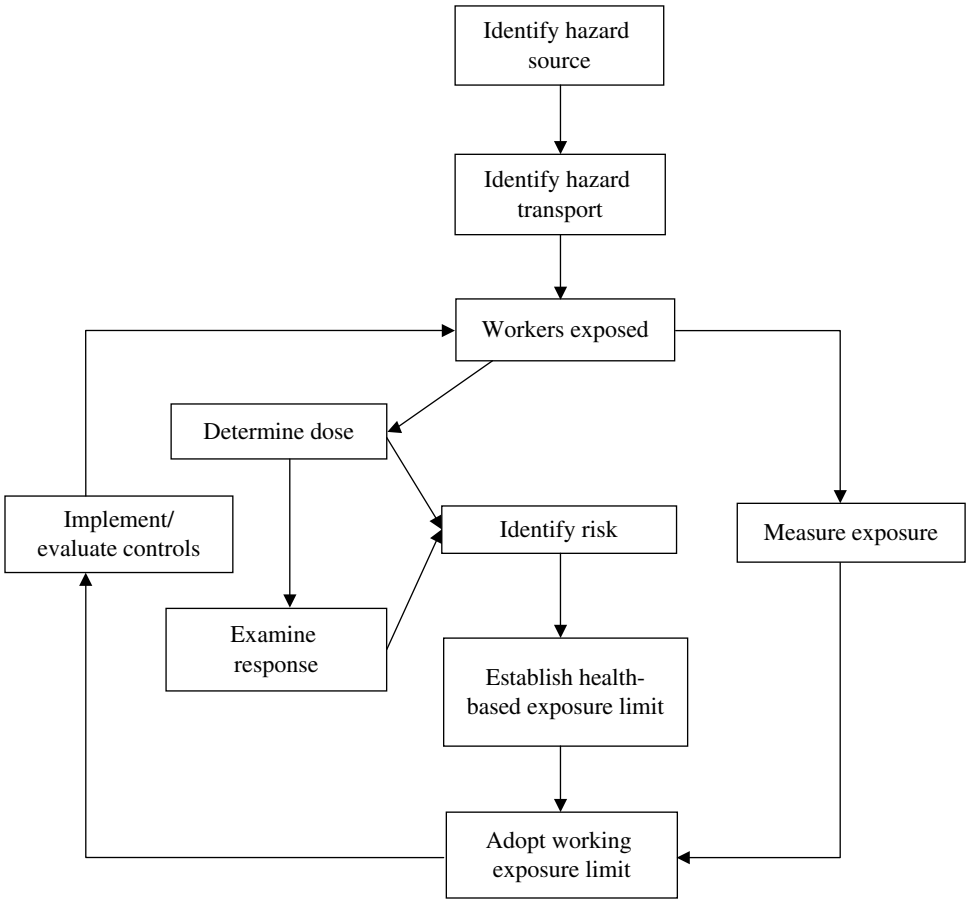


FIGURE 22.1 Process for assessing occupational health hazards and risks.

TABLE 22.4 Evaluation of Health Risk from Dust Exposure

Step	Possible Task	OH Professional	Collaborators/Comments
Identify hazard source	Identify composition, size, and source of dust	Industrial hygiene/safety professional	Engineering/production manager
Identify hazard transport	Determine how far dust transports	Industrial hygiene	—
Workers exposed	Identify workers and group by relative exposure	Industrial hygiene/safety professional/nurse	Human resources/area supervisor
Determine dose	Determine inhaled dose and monitor exposures	Industrial hygiene	—
Examine response	Examine literature and use medical diagnostics	Industrial hygiene/nurse/physician	Consult toxicologist if literature is lacking
Identify risk	Identify health risks for range of exposures studied in dose/response	Industrial hygiene	Toxicologist, epidemiologist, physician
Establish health-based exposure limit	Incorporate dose–response data with available exposure limits and documentation	Industrial hygiene	Engineering—feasibility input review by physician
Adopt working exposure limit	Finalize facility-specific exposure limit for this source	Industrial hygiene	Management needed to accept
Implement/evaluate controls	Improve ventilation system	Industrial hygiene/safety	Engineering/facilities
	Institute respiratory protection (RP) program		Physician/nurse—health evaluation for fitness to wear RP
Continued exposure assessment	Scheduled monitoring of exposure and workers' health	Industrial hygiene/safety physician/nurse	Annually if RP is required

TABLE 22.5 Evaluation of Health Risk from Lifting Task

Step	Task	Primary OH Professional	Additional Collaborators
Identify hazard source	Identify at-risk lifting tasks	Safety/IH/nurse	Ergonomist
Identify hazard transport	Identify additional full-shift lifting tasks	Safety/IH/nurse	Area supervisor
Workers exposed	Identify workers and groups by relative exposure	Safety/IH/nurse	Area supervisor
Determine dose	Determine exposure forces and patterns	Ergonomist	Nurse
Examine response	Evaluate MSD cases in exposed group	Nurse/physician	Ergonomist
Identify risk	Identify health risks for range of exposures studied in dose/response	Ergonomist	
Establish health-based exposure limit	Recommend maximum lift/exposure duration	Ergonomist	Safety
Adopt working exposure limit	Finalize facility-specific lift/exposure duration	Safety manager	Nurse
Implement/evaluate controls	Reengineer the job or work schedule to reduce exposures	Ergonomist, safety manager	Engineering, area supervisor, affected workers
Continued exposure assessment	Review ongoing implementation of work practice/job rotation effectiveness	Safety, nurse	

recommendations for exposure limits to OSHA. The NIOSH is also required to support the education and training of occupational health professionals and to fund research by academic investigators.

The U.S. Environmental Protection Agency (EPA) was created to protect human health and the environment by developing and enforcing environmental regulations. When Congress created the EPA in 1970, it also passed the Clean Air Act. This was followed in 1972 by the Clean Water Act, which was created to prevent pollution of the nation's surface waters. Other major environmental regulations followed, including the Toxic Substances Control Act (TSCA) in 1976, the Resource Conservation and Recovery Act (RCRA) of 1976, and the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) of 1980 (more commonly known as the *Superfund*). The latter two of these acts focus on the tracking and management of hazardous waste materials. The TSCA also gives the EPA

the authority to require a manufacturer/importer of new chemicals to the United States to (i) submit a premanufacture notice that triggers the EPA's Office of Pollution Prevention and Toxics to perform a risk assessment for the new chemical and (ii) requires the manufacturer to protect employees below a new chemical exposure limit (NCEL) that is specified by the EPA from the risk assessment. The EPA also has regulatory authority to provide guidelines on indoor air quality, with regional experts available to provide guidance on radon, mold and moisture, and general indoor air quality concerns.

22.9 ETHICAL CONSIDERATIONS FOR THE OCCUPATIONAL AND ENVIRONMENTAL HEALTH PROFESSIONAL

Occupational and environmental health providers, regardless of discipline, have at least one professional organization dedicated to supporting their continued education and certification. Because of the potential for conflicts of interest and ambiguity of obligation, each of these professional organizations has a code of ethics that specify the scope of work and ethical behavior to which members must adhere. While specific details for each professional organization are provided, it is important to note that there are some common elements to each of these professional codes. Specifically, each requires professionals act to protect worker confidentiality, provide services within one's scope of expertise, represent all facts relevant to health and safety, and engage in ongoing professional development and education. Because occupational health providers often "serve two masters" (the workers they protect and the company that employs them), published codes of ethics provide an important framework for promoting professional practices that benefit those at risk of harm from hazardous exposures, even if not encouraged to do so by employers. Details are included in each of the discussions that follow.

Occupational Medicine

Just as with other fields of healthcare, ethical principles have been developed to guide occupational and environmental health providers. The first ethical principle is *confidentiality*. This principle is a cornerstone of the physician–patient relationship upheld throughout all fields of medicine. It is, however, particularly challenging in the field of occupational and environmental medicine. Confidentiality applies to both the written and verbal release of protected health information. In the United States, the *Health Insurance Portability and Accountability Act* (HIPAA) establishes the conditions under which healthcare providers may share a patient's protected health information. As such, medical records need to be maintained in a protected manner, and the sharing of patients' protected health information, in general, can only occur if the patient provides written consent to release the information. Confidentiality is not absolute, however. In certain legal settings, for example, a request for workers' compensation benefits, the employer and legal agencies involved in the adjudication of the claim have the right to access the employee's health information, *with regard to the claim*. In working with employers and workers' compensation insurance companies, occupational medicine providers are required to inform the workers that health information collected during the examination is not confidential and will be released to the employer and other third parties in order to process the claim.

In some cases, the duty to maintain confidentiality may conflict with other ethical duties, in particular, those of *nonmaleficence* and *beneficence*. Nonmaleficence is best thought of as the ancient requirement for healthcare providers to "do no harm." Beneficence goes beyond this requirement and obliges professionals to use their authority to be of benefit to patients and others. If, for example, during a physician–patient interaction, it becomes clear that other workers may also be at risk of adverse health effects by the same exposure that caused the patient to seek medical attention, the physician may have an ethical (and sometimes legal) responsibility to report this to third parties such as employers or public health departments. In these cases, the duty to protect the confidentiality of the patient may be in conflict with the need to protect the health of other exposed workers. Legal obligations may also be relevant in such situations. When ethical obligations are in conflict, it is critical for the practitioner to carefully examine the consequences of each possible action and to act in the manner that is most protective of the health of the public and most respectful of the rights of patients and workers.

Although it may seem obvious, an occupational medicine physician must always inform patients about their medical conditions, including discussions with the patient on the diagnostic and treatment options. This is also no different than any other physician–patient interaction. Given the setting, however, the provider also needs to disclose to patients therapeutic recommendations that may include relocation or job modification to avoid continued reexposure to certain hazardous environments.

The American College of Occupational and Environmental Medicine (ACOEM), the professional organization for occupational and environmental providers, has a code of ethical conduct to guide practitioners in the field (ACOEM 2012).

Occupational Health Nursing

The American Association of Occupational Health Nurses (AAOHN) has a code of ethics for OHNs. The code requires that OHNs provide health, wellness, safety, and other

related services to clients with regard for human dignity and rights, unrestricted by consideration of social or economic status, personal attributes, or the nature of the health status. In addition, the AAOHN, like the ACOEM standard described earlier, requires the protection of confidential information and authorizes release only as "required or permitted by law." OHNs are also required to maintain competence and accept responsibility for professional judgment and actions.

Industrial Hygiene

The ABIH established the code of ethics for CIHs. This code requires that hygienists maintain high standards of integrity and professional conduct, accept responsibility for their actions, and continually seek to enhance their professional capabilities. Specific responsibilities to clients, employers, employees, and the public are also defined. First and foremost, the code requires that industrial hygiene services be delivered with objective and independent professional judgment. Because of the diverse nature of expertise in the comprehensive industrial hygiene field, professionals are required to recognize their limitations and only provide services for which they are qualified.

Industrial hygienists are required to maintain and respect confidentiality of sensitive information, as are other occupational health professionals. However, the industrial hygiene code of ethics clearly defines the obligation of CIHs to release information if directed to by the courts or government agency or if failure to release the information would likely result in death or serious physical harm to employees or the public.

Additional ethical obligations require the hygienist to properly use credentials, to only use the CIH designation for work that is under his/her control, to provide truthful and accurate presentations in public statements, and to recognize the intellectual property rights of others. The ABIH frequently revises its code of ethics and announces changes to their membership. The full code of ethics is available at the ABIH website.

Ergonomists

The CCCPE code of ethics outlines professional responsibilities of confidentiality, recordkeeping, integrity, and impartiality. Specific concepts of integrity are outlined, including disclosing conflicts of interest; adequately representing facts and expressing opinions, with clear distinction; informing clients of limitations in his own qualifications and in the work product or services provided; and immediately informing the client of any error that may have been made. Similar to the industrial hygiene code, an ergonomist may not claim skills beyond his capability and is obligated to act in the interest of the client.

22.10 SUMMARY AND CONCLUSION

Despite regulations and recommended standards put in place to protect workers' health, serious hazards continue to exist in occupational settings. Identifying hazards in the work environment requires the expertise of many occupational and environmental health providers and is most efficient and effective when these professionals work as a team. While occupational hazards change over time and new hazards emerge with the development of new technologies and production techniques, the methods applied to assessing and controlling hazards to evaluating and treating affected workers are universal. The dedication and commitment of occupational health practitioners, relying on scientific data for determining risk and making evidence-based interpretations, are essential to the protection of workers' health.

The main points of this chapter are:

- Occupational injuries and illnesses are extremely common and include a wide spectrum of conditions.
- The historical foundations of occupational and environmental health date back to the 1700s and have been shaped by ongoing advances in technology and industry.
- The three overarching goals of occupational and environmental health are primary (prevent exposure), secondary (prevent disease development), and tertiary (prevent disease from progressing) prevention of injury and illness.
- Occupational diseases are often indistinguishable from nonoccupational disease and are multifactorial in nature.
- Occupational health providers perform many activities including diagnosis and treatment of disease, evaluation of fitness for duty, assessment of causality, medical surveillance, and education of workers and employers.
- The occupational health field is composed of practitioners from many disciplines who work together to optimize the health and safety of workers.
- Practitioners in occupational and environmental health must adhere to a widely accepted code of ethics.

ACKNOWLEDGMENT

The authors are grateful to Drs Howard Frumkin and Edward Galaid for their contribution to the chapter on this topic in the previous edition of this book.

SUGGESTED READING

Abrams HK. A short history of occupational health. J Public Health Policy 2001;22(1):34–80.

American Association of Occupational Health Nurses. 2009. Code of Ethics and Interpretive Statements. Available at http://www.aaohn.org/membership/membership-marketplace/code-of-ethics-and-interpretive-statements-details.html. Accessed July 11, 2014.

American College of Occupational and Environmental Medicine. 2001. Careers in Occupational and Environmental Medicine. Available at http://www.acoem.org/Page2Column.aspx?PageID=7392&id=992. Accessed March 13, 2011.

American College of Occupational and Environmental Medicine. 2011. ACOEM Code of Ethics. Available at http://www.acoem.org/codeofconduct.aspx. Accessed July 11, 2014.

American Industrial Hygiene Association. 2007. Joint Industrial Hygiene Associations Member Ethical Principles. Available at https://www.acgih.org/Member_Ethical_Principles.pdf. Accessed July 11, 2014.

Blanc P. Occupational and environmental medicine: the historical perspective. In: Rosenstock L, Cullen MR, Brodkin CA, Redlich CA, editors. *Textbook of Clinical Occupational and Environmental Medicine*. 2nd ed. Philadelphia: Elsevier Saunders; 2005.

Brodkin CA, Rosenstock L. Asbestosis and asbestos-related pleural disease. In: Rosenstock L, Cullen MR, Brodkin CA, Redlich CA, editors. *Textbook of Clinical Occupational and Environmental Medicine*. 2nd ed. Philadelphia: Elsevier Saunders. 2005

Canadian Centre for Occupational Health and Safety. 2006. Hazard Control. OSHA Answers. Available at http://www.ccohs.ca/oshanswers/hsprograms/hazard_control.html. Accessed August 7, 2014.

Centers for Disease Control and Prevention (CDC) 2011. The National Institute of Occupational Safety and Health (NIOSH). Available at http://www.cdc.gov/niosh/about.html. Accessed April 15, 2011.

Cullen MR, Rosenstock L. Approach to the patient. In: Rosenstock L, Cullen MR, Brodkin CA, Redlich CA, editors. *Textbook of Clinical Occupational and Environmental Medicine*. 2nd ed. Philadelphia: Elsevier Saunders; 2005a.

Cullen MR, Rosenstock L. Social issues in the practice of occupational and environmental medicine. In: Rosenstock L, Cullen MR, Brodkin CA, Redlich CA, editors. *Textbook of Clinical Occupational and Environmental Medicine*. 2nd ed. Philadelphia: Elsevier Saunders; 2005b.

Cullen MR, Rosenstock L, Kilbourne EM. Introduction to occupational and environmental medicine. In: Rosenstock L, Cullen MR, Brodkin CA, Redlich CA, editors. *Textbook of Clinical Occupational and Environmental Medicine*. 2nd ed. Philadelphia: Elsevier Saunders; 2005.

Franco G. Ramazzini and Worker's Health. Lancet 1999;354; 858–861.

Hamilton A. Lead poisoning in the United States. Am J Public Health 1914;4:477–480.

Hammond EC, Selikoff IJ, Seidman H. Asbestos exposure, cigarette smoking and death rates. Ann N Y Acad Sci 1979;330: 473–490.

Herzstein J. Susceptible populations. In: Rosenstock L, Cullen MR, Brodkin CA, Redlich CA, editors. *Textbook of Clinical Occupational and Environmental Medicine*. 2nd ed. Philadelphia: Elsevier Saunders; 2005.

Keyserling W. Monroe Occupational ergonomics: promoting safety and health through work designs. In: Levy BS, Wegman DH, Baron SL, Sokas RK, editors. *Occupational and Environmental Health: Recognizing and Preventing Disease and Injury*. 5th ed. Philadelphia: Lippincott Williams & Wilkins; 2006.

Levy BS, Wegman DH, Baron SL, Sokas RK. Occupational and environmental health: an overview. In: Levy BS, Wegman DH, Baron SL, Sokas RK, editors. *Occupational and Environmental Health: Recognizing and Preventing Disease and Injury*. 5th ed. Philadelphia: Lippincott Williams & Wilkins; 2006.

Lukes E. The nursing process and program planning. AAOHN J 2010;58(1):5–7.

Occupational Safety and Health Administration. About OSHA. Available at http://www.osha.gov/about.html. Accessed April 15, 2011.

Occupational Safety and Health Administration (OSHA). OSHA Lead Standard. Available at http://www.osha.gov/pls/oshaweb/owadisp.show_document?p_table=STANDARDS&p_id=10030. Accessed April 11, 2011.

Occupational Safety and Health Administration. Substance Safety Data Sheet. Available at http://www.osha.gov/pls/oshaweb/owadisp.show_document?p_id=10036&p_table=STANDARDS. Accessed January 15, 2012.

Parker JE. The occupational and environmental history and examination. In: Rom W, editor. *Environmental and Occupational Medicine*. 4th ed. Philadelphia: Wolters Kluwer/Lippincott Williams & Wilkins; 2007.

Rom WN. The discipline of environmental and occupational medicine. In: Rom W, editor. *Environmental and Occupational Medicine*, 4th ed. Philadelphia: Wolters Kluwer/Lippincott Williams & Wilkins; 2007.

Silverstein B, Watters MT. Regulations in practice: assessing and enforcing compliance with health and safety regulations. In: Levy BS, Wegman DH, Baron SL, Sokas RK, editors. *Occupational and Environmental Health: Recognizing and Preventing Disease and Injury*. 5th ed. Philadelphia: Lippincott Williams & Wilkins; 2006.

Smith TJ, Schneider T. Occupational and environmental hygiene. In: Levy BS, Wegman DH, Baron SL, Sokas RK, editors. *Occupational and Environmental Health: Recognizing and Preventing Disease and Injury*. 5th ed. Philadelphia: Lippincott Williams & Wilkins; 2006.

Tarlo SM, Chan-Yeung M. Occupational asthma. In: Rosenstock L, Cullen MR, Brodkin CA, Redlich CA, editors. *Textbook of Clinical Occupational and Environmental Medicine*. 2nd ed. Philadelphia: Elsevier Saunders; 2005.

U.S. Department of Labor, Bureau of Labor Statistics. 2010. Nonfatal Occupational Injuries and Illnesses Requiring Days Away from Work, 2009. USDL-10-1546. Available at http://www.bls.gov/news.release/archives/osh2_11092010.pdf. Accessed July 11, 2014.

23

HUMAN HEALTH RISK ASSESSMENT

LEAH D. STUCHAL, ROBERT C. JAMES, AND STEPHEN M. ROBERTS

Risk assessment is an ever-evolving process whereby scientific information on the hazardous properties of chemicals and the extent of exposure results in a statement as to the probability that exposed populations will be harmed. The probability of harm can be expressed either qualitatively or quantitatively, depending on the nature of the scientific information available and the intent of the risk assessment. Risk assessment is not research per se, but rather a process of collecting and evaluating existing data. As such, risk assessment draws heavily on the disciplines of toxicology, epidemiology, pathology, molecular biology, biochemistry, mathematical modeling, industrial hygiene, analytical chemistry, and biostatistics. The certainty with which risks can be accurately assessed, therefore, depends on the conduct and publication of basic and applied research relevant to risk issues. While firmly based on scientific considerations, risk assessment is often an uncertain process requiring considerable judgment and assumptions on the part of the risk assessor. Ultimately, the results of risk assessments are integrated with information on the consequences of various regulatory options in order to make decisions about the need for, method of, and extent of risk reduction.

It is clear that society is willing to accept some risks in exchange for the benefits and conveniences afforded by chemical use. After all, we knowingly apply pesticides to increase food yield, drive pollutant-emitting automobiles, and generate radioactive wastes in the maintenance of our national defense. We legally discharge the by-products of manufacturing into the air we breathe, the water we drink, and the land on which our children play. In addition, we have a history of improper waste disposal, the legacy of which is thousands of uncontrolled hazardous waste sites. To ensure that the risks posed by such activities are not unacceptably large, it is necessary to determine safe exposure levels in the workplace and environment. Decisions must also be made on where to locate industrial complexes, on remediation options for hazardous waste sites, tolerance levels for pesticides in foods, safe drinking water standards, air pollution limits, and the use of one chemical in favor of another. Risk assessment provides the tools to make such determinations.

This chapter provides an overview of the risk assessment process and discusses:

- The basic steps of risk assessment
- How risk assessments are performed in a regulatory context
- Differences between human health and ecological risk assessments
- Differences in the estimation of cancer and noncancer risks
- Differences between deterministic and probabilistic risk assessments
- Issues associated with estimating risks from chemical mixtures
- Comparisons of risks from chemical exposure with other health risks
- Risk communication from chemical exposure with other health risks

Principles of Toxicology: Environmental and Industrial Applications, Third Edition. Edited by Stephen M. Roberts, Robert C. James and Phillip L. Williams.

23.1 RISK ASSESSMENT BASICS

A Basic Risk Assessment Paradigm

In 1983, the National Research Council described risk assessment as a four-step analytical process consisting of hazard identification, dose–response assessment, exposure assessment, and risk characterization. These fundamental steps have achieved a measure of universal acceptance and provide a logical framework to assemble information on the situation of potential concern and provide risk information to inform decision making (Figure 23.1). The process is rigid enough to provide some methodological consistency that promotes the reliability, utility, and credibility of risk assessment outcomes, while at the same time allowing for flexibility and judgment by the risk assessor to address an endless variety of risk scenarios. Each step in the four-step process known as *risk assessment* is briefly discussed below.

Step 1: Hazard Identification. The process of determining whether exposure to a chemical agent, under any exposure condition, can cause an increase in the incidence or severity of an adverse health effect (cancer, birth defect, neurotoxicity, etc.). Although the matter of whether a chemical can, under any exposure condition, cause cancer or other adverse health effect is theoretically a yes/no question; there are few chemicals for which the human data are definitive. Therefore, laboratory animal studies, *in vitro* tests, and structural and mechanistic comparability to other known chemical hazards are considered in addition to the epidemiological data. This step is common to qualitative and quantitative risk assessment.

Step 2: Dose–Response Assessment. The process of characterizing the relationship between the dose of a chemical and the incidence or severity of an adverse health effect in the exposed population. A dose–response assessment factors not only in the magnitude, duration, and frequency of exposure but also other potential response-modifying variables such as age, sex, and certain lifestyle factors. A dose–response assessment frequently requires extrapolation from high to low doses and from animals to humans.

Step 3: Exposure Assessment. The process of specifying the exposed population, identifying potential exposure routes, and measuring or estimating the magnitude, duration, and frequency of exposure. Exposure can be assessed by direct measurement or estimated with a variety of exposure models. Exposure assessment can be quite complex because exposure frequently occurs

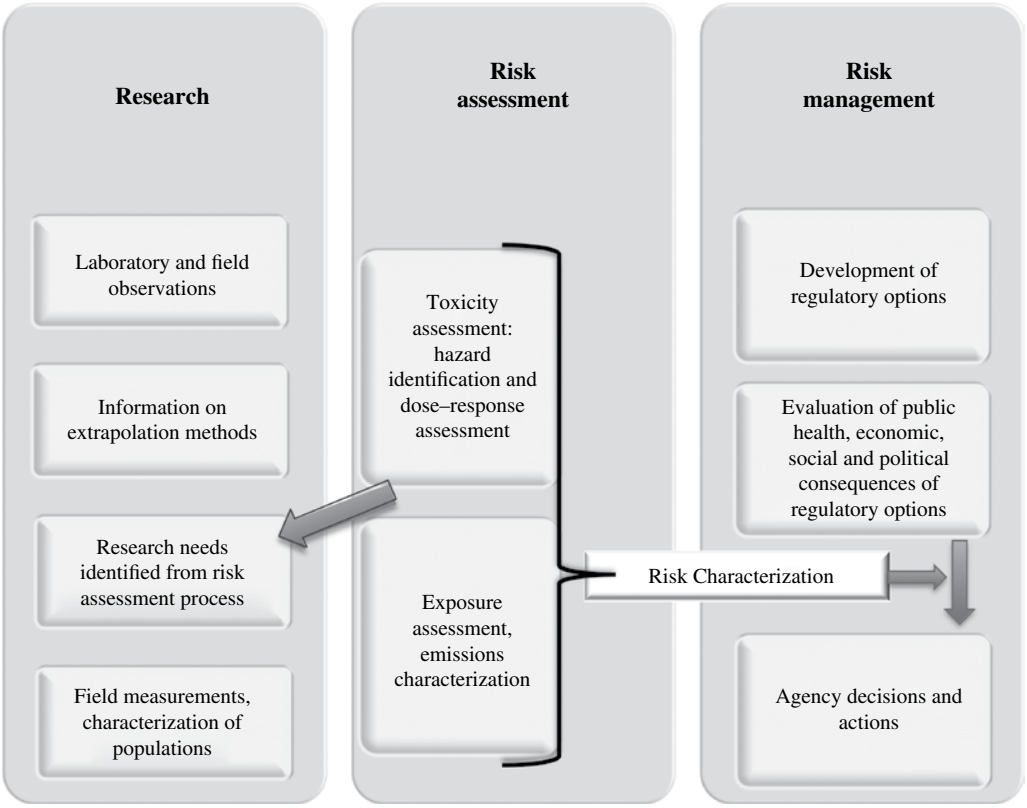


FIGURE 23.1 Elements of risk assessment and risk management. Risk assessment provides a means to organize and interpret research data in order to inform decisions regarding human and environmental health. Through the risk assessment process, important data gaps and research needs are often identified, assisting in the prioritization of basic and applied toxicological research. *Source:* Adapted from NRC (1983).

to a mixture of chemicals from a variety of sources (air, water, soil, food, etc.).

Step 4: Risk Characterization. The integration of information from steps 1 to 3 to develop a qualitative or quantitative estimate of the likelihood that any of the hazards associated with the chemical(s) of concern will be realized. The characterization of risk must often encompass multiple populations having varying exposures and sensitivities. This step is particularly challenging as a variety of data must be assimilated and communicated in such a way as to be useful to everyone with an interest in the outcome of the risk assessment. This may include not only governmental and industry risk managers but also the public as well. This step includes a descriptive characterization of the nature, severity, and route dependency of any potential health effects, as well as variation within the population(s) of concern. Any uncertainties and limitations in the analysis are described in the risk characterization, so that the strengths, weaknesses, and overall confidence in the risk estimates can be understood.

Circumstances may exist in which no risk can be inferred from an exposure assessment that reveals no opportunity for individuals to receive a dose of the chemical. Therefore, situations sometimes exist where a comprehensive risk assessment is unnecessary. In such instances, it may be more practical to communicate findings in a qualitative manner, that is, to state it is highly unlikely chemical X will pose any significant health risk because there is no exposure to the chemical. At other times, quantitative expressions of risk might be more appropriate, as in the case of a population chronically exposed to a known human carcinogen in drinking water. An expression of such risk might be that the lifetime excess cancer risk from exposure is 3 in 1,000,000 (or 3×10^{-6}). Often, such numerical expressions of risk convey an unwarranted sense of precision by failing to communicate the uncertainty inherent in their derivation. They may also prove difficult for nontechnical audiences to comprehend. On the other hand, qualitative risk estimates may appear more subjective and not invoke the same degree of confidence in the risk assessment findings as a numerical expression of risk. Also, qualitative expressions of risk do not readily allow for comparative risk analyses, a useful exercise for putting added risk into context. Although addressed later in this chapter, it is worth mentioning here that effective risk communication plays a key role in utilizing risk assessment findings for the protection of public health.

Risk Assessment in a Regulatory Context: The Issue of Conservatism

Regulatory agencies charged with protecting public health and the environment are constantly faced with the challenge of setting permissible levels of chemicals in the home, workplace, and natural environment. For example, the Occupational Safety and Health Administration (OSHA) is responsible for setting limits on chemical exposure in the workplace, the Food and Drug Administration (FDA) has permissible limits on chemicals such as pesticides in the food supply, and the Environmental Protection Agency (U.S. EPA) regulates chemical levels in air, water, and sometimes soil. Ideally, the level of chemical contamination or residues in many of these media (food, water, air, etc.) would be zero, but this simply is not feasible in a modern industrial society. Although it may not be possible to completely eliminate the presence of unwanted chemicals from the environment, there is almost universal agreement that we should limit exposures to these chemicals to levels that do not cause illness or environmental destruction. The process by which regulatory agencies set limits with this goal in mind is a combination of risk assessment and risk management.

The risks associated with chemical exposure are not easily measured. While studies of worker health have been extremely valuable in assessing risks and setting standards for occupational chemical exposure, determining risks from lower doses typically associated with environmental exposures has been difficult. Epidemiologic studies of environmental chemical exposure can provide some estimate of increased risk of specific diseases associated with a particular chemical exposure compared with a control population, but there are several problems in attempting to generalize the results of such studies. Exposure of a population is often difficult to quantify, and the extrapolation of observations from one situation to another (e.g., different populations, different manners of exposure, different exposure levels, different exposure durations) is challenging. For the most part, risk assessments for environmental chemical exposures must rely on modeling and assumptions to generate estimates of potential risks. Because these risk estimates usually cannot be verified, they represent hypothetical or theoretical risks. This is an important facet of risk assessment that is often misunderstood by those who erroneously assume that risk estimates for environmental chemical exposure have a strong empirical basis.

As discussed in subsequent sections, there are many sources of uncertainty in deriving risk estimates. Good data regarding chemical exposure and uptake are seldom available, forcing reliance on models and assumptions that may or may not be valid. Toxicity information often must be extrapolated from one species to another (e.g., use of data from laboratory mice or rats for human health risk assessment), from one route of exposure to another (e.g., use of toxicity data following ingestion to evaluate risks from dermal exposure), and from high doses to the lower doses more commonly encountered with environmental exposure. In view of all of these uncertainties, it is impossible to develop precise estimates of risks from chemical exposures. Choices made by the risk assessor, such as

which exposure model to use or how to scale doses when extrapolating from rodents to humans, can have a profound impact on the risk estimate.

Regulatory agencies address uncertainty in risk assessments by using conservative approaches and assumptions; that is, in the face of scientific uncertainty, they will select models and assumptions that tend to overestimate, rather than underestimate, risk so as to be health protective. Since most risk assessments are by, or for, regulatory agencies, this conservatism is a dominant theme in risk assessments and a continuous source of controversy. Some view the conservatism employed by regulatory agencies as excessive, resulting in gross overestimation of risks and unwarranted regulations that waste billions of dollars. Others question whether regulatory agencies are conservative enough and suggest that the public (particularly more sensitive individuals such as children) may not be adequately protected by contemporary risk assessment approaches.

Defining Risk Assessment Problems

A coherent risk assessment requires a clear statement of the risk problem to be addressed. This should be developed very early in the risk assessment process and is shaped by the question(s) the risk assessment is expected to answer. Ideally, both the risk assessor(s) and the individuals or organizations that will ultimately use the risk assessment will have input. This helps ensure that the analysis will be technically sound and serve its intended purpose.

One of the first issues to address is which chemicals or agents should be included in the analysis. In some situations, this may be straightforward, such as a risk assessment focused specifically on occupational exposure to a particular chemical. In other circumstances, the chemicals of concern may not be obvious. An example of this would be risk assessment for a chemical disposal site where the chemicals present and their amounts are initially unknown. A related issue is which health effects the risk assessment should address. While it is tempting to answer "all of them," it must be recognized that each chemical in a risk assessment is capable of producing a variety of adverse health effects, and the dose–response relationships for these effects can vary substantially. Developing estimates of risks for each of the possible adverse effects of each chemical of interest is usually impractical. A simpler approach is to estimate risks for the health effect to which individuals are most sensitive, specifically, the one that occurs at the lowest dose. If individuals can be protected from this effect, whatever it might be, they will logically be protected from all other effects. Of course, this approach presumes that the most sensitive effect has been identified and dose–response relationship information for this effect exists. Obviously, for this approach to be effective, the toxicology of each chemical of interest must be reasonably well characterized.

In defining the risk problem, populations potentially at risk must be identified. These populations would be groups of individuals with distinct differences in exposure, sensitivity to toxicity, or both. For example, a risk assessment for a contaminated site might include consideration of workers at the site, occasional trespassers or visitors to the site, or individuals who live at the site if the land is (or might become) used for residential purposes. If residential land use is contemplated, risks are often calculated separately for children and adults, since they may be exposed to different extents and therefore have different risks. Depending on the goals of the risk assessment, risks may be calculated for one or several populations of interest.

Many chemicals move readily in the environment from one medium to another. Thus, a chemical spilled on the ground can volatilize into the air, migrate to groundwater and contaminate a drinking water supply, or be carried with surface water runoff to a nearby stream or lake. Risk assessments have to be cognizant of environmental movement of chemicals and the fact that an individual can be exposed to chemicals by a variety of pathways. In formulating the risk problem, the risk assessor must determine which of many possible pathways are complete; that is, which pathways will result in movement of chemicals to a point where contact with an individual will occur. Each complete pathway provides the opportunity for the individual to receive a dose of the chemical and should be considered in some fashion in the risk assessment. Incomplete exposure pathways—those that do not result in an individual coming in contact with contaminated environmental media (e.g., air, water, soil)—can be ignored because they offer no possibility of receiving a dose of chemical and therefore pose no risk.

Risk assessments can vary considerably in the extent to which information on environmental fate of contaminants is included in the analysis. Some risk assessments, for example, have attempted to address risks posed by chemicals released to the air in incinerator emissions. These chemicals are subsequently deposited on the ground where they are taken up by forage crops that are consumed by dairy cattle. Consumption of meat or milk from these cattle is regarded as a complete exposure pathway from the incinerator to a human receptor. As the thoroughness of the risk assessment increases, so does the complexity. As a practical matter, complete exposure pathways that are thought to be minor contributors to total exposure and risk are often acknowledged but not included in the calculation of risk to make the analysis more manageable.

Often, exposure can lead to uptake of a chemical by more than one route. For example, contaminants in soil can enter the body through dermal absorption, accidental ingestion of small amounts of soil, or inhalation of contaminants volatilized from soil or adherent to small dust particles. Consequently, the manner of anticipated exposure is important to consider, as it will dictate the routes of exposure (i.e., inhalation, dermal contact, or ingestion) that need to be included in the risk assessment for each exposure scenario.

Human Health versus Ecological Risk Assessments: Fundamental Differences

Ecological risk assessments are defined as those that address species other than humans, namely, plant and wildlife populations. Problem formulation is more challenging when conducting ecological risk assessments. Instead of one species, there are several to consider. Also, the exposure pathway analysis is more complicated, at least in part because some of the species of interest consume other species of interest, thereby acquiring their body burden of chemical. Unlike human health risk assessments, where protection of individuals against any serious health impact is nearly always the objective, goals for ecological risk assessments are often at the population, or even ecosystem, level rather than focusing on individual plants and animals. Consequently, development of assessment and measurement endpoints consistent with the goals of the ecological risk assessment is essential in problem formulation for these kinds of analyses.

Historically, the risk assessment process has focused primarily on addressing potential adverse effects to exposed human populations, and the development of well-defined methods for human health risk assessment preceded those for ecological risk assessment. However, increasing concern for ecological impacts of chemical contamination has led to a "catching up" in risk assessment methodology. While detailed methods for both human health and ecological risk assessment are now in place, they are not identical. The conceptual basis may be similar, including some form of hazard identification, exposure assessment, dose–response assessment, and risk characterization. However, there are some important differences in approaches, reflecting the reality that there are some important differences in evaluating potential chemical effects in humans versus plants and wildlife.

The most obvious difference between human health and ecological risk assessments is that the ecological risk assessments are inherently more complicated. Human health risk assessments, of course, deal with only one species. Ecological risk assessments can involve numerous species, many of which may be interdependent. Given the nearly endless array of species of plants and animals that might conceivably be affected by chemical exposure, there must be some process to focus on species that are of greatest interest to keep the analysis to a manageable size. A species may warrant inclusion in the analysis because it is threatened or endangered, because it is a species on which many others depend (e.g., as a food source), or because it is especially sensitive to toxic effects of the chemical and can therefore serve as a sentinel for effects on other species.

The increased complexity of analysis for ecological risk assessments extends to evaluation of exposure. In human health risk assessment, the potential pathways by which the chemical(s) of interest can reach individuals must be assessed and, if possible, the doses of chemicals received by these pathways estimated. In an ecological risk assessment, the same process must be undertaken, but for several species instead of just one. Also, an ecological risk assessment typically must evaluate food chain exposure. This is particularly important when chemicals of interest tend to bioaccumulate, resulting in very high body burdens in predator species at the top of the food chain. Not only must the potential for bioaccumulation be assessed, but also the escalating doses for species of interest must be estimated according to their position in the food chain. This type of analysis is only included in human health risk assessments when estimating dose from a potential food source (e.g., fish, meat, or milk).

A third distinction between human health and ecological risk assessment lies in the assessment objectives. Human health risk assessments characteristically focus on the most sensitive potential adverse health effect, specifically, that which occurs at the lowest dose. In this way, they are directed to evaluating the potential for *any* health effect to occur. For ecological risk assessments, the analyses generally address only endpoints that affect fecundity (growth, survival, and reproduction). Thus, the goal of an ecological risk assessment might be to determine whether the presence of a chemical in the environment at a particular concentration would result in declining populations for a specific species (e.g., due to mortality or reproductive failure), disappearance of a species in a particular area, or loss of an entire ecosystem, depending on risk management objectives. It is entirely possible that chemical exposure could result in the deaths of many animals, but as long as the populations were stable, the risk would be considered acceptable. The exception to population protection is for species with special legal protection (endangered, threatened, or listed). These species should be protected on an individual level, and all adverse effects should be considered in determining a critical effect.

What constitutes an unacceptable impact is not clearly defined in ecological risk assessment. Regulatory agencies may decide to focus on higher tropic level species or may not focus on a species at all, protecting the habitat instead. Alternatively, biodiversity may be utilized as an ecological endpoint. These alternative endpoints allow for the loss of entire populations and for the establishment of nonnative and invasive species. This reflects philosophical and risk management differences in terms of what constitutes an unacceptable chemical impact on humans versus plants or wildlife.

Because of the greater potential complexity of an ecological risk assessment, more attention must be given to ensuring that an analysis of appropriate scope and manageable size is achieved. For this reason, ecological risk assessments are more iterative in nature than their human health counterparts. An ecological risk assessment begins with a screening-level assessment, which is a form of preliminary investigation to determine whether unacceptable risks to ecological receptors may exist. It includes a review of data regarding chemicals present and their concentrations, species present, and potential

pathways of exposure. It is a rather simplified analysis that uses conservative or worst-case assumptions regarding exposure and toxicity. If the screening analysis finds no indication of significant risks using very conservative models and assumptions, the analysis is concluded. If the results of the screening analysis suggest possible ecological impacts, a more thorough analysis is conducted that might include additional samples of environmental media, taking samples of wildlife to test for body burdens of chemicals, carefully assessing the health status of populations exposed to the chemical, conducting toxicity tests, conducting more sophisticated fate and transport analysis of the chemicals of potential concern, and a more detailed and accurate exposure assessment.

23.2 HAZARD IDENTIFICATION

Hazard identification involves an assessment of the intrinsic toxicity of the chemical(s) of potential concern. This assessment attempts to identify health effects characteristically produced by the chemical(s) that may be relevant to the risk assessment. While this may appear to be a straightforward exercise, in reality, it requires a good deal of careful analysis and scientific judgment. The reason for this is that the risk assessor rarely has the luxury of information that adequately describes the toxicity of a chemical under the precise set of circumstances to be addressed in the risk assessment. Instead, the risk assessor typically must rely on incomplete data derived from species other than the one of interest under exposure circumstances very different from those being evaluated in the risk assessment. The existence in the scientific literature of poorly designed studies with misleading results and conclusions, as well as conflicting data from seemingly sound studies, further complicates the task.

This section of the chapter discusses some of the considerations when reviewing and evaluating the toxicological literature for assessment of intrinsic toxicity. Many of these considerations address suitability of data for extrapolation from one set of circumstances to another, while others pertain to the fundamental reliability of the information. Much of the discussion regarding extrapolation deals with assessing the value of animal data in predicting responses in humans, since human health risk assessments are forced to rely predominantly on animal studies for toxicity data. Keep in mind that most of the same extrapolation issues are equally relevant for ecological risk assessments, where often toxicity in wildlife species has to be inferred from data available only from laboratory animal species.

Information from Epidemiologic Studies and Case Reports

Observations of toxicity in humans can be extremely valuable in hazard identification. They offer the opportunity to test the applicability of observations made in animal studies to humans and may even provide an indication of the relative potency of the chemical in humans versus laboratory animal models. If the human studies are of sufficient size and quality, they may stand alone as the basis for hazard identification in human health risk assessment.

Despite the attractiveness of human studies, they often have significant limitations. A less-than-rigorous effort to properly match exposed and control populations makes it difficult or impossible to attribute observed differences in health effects to chemical exposure with any confidence. Even in well-designed epidemiologic studies, there is always the possibility that an unknown critical factor causally related to the health effect of interest has been missed. For this reason, a consistent association between chemical exposure and a particular effect in several studies is important in establishing whether the chemical produces that effect in humans.

Other criteria in evaluating epidemiologic studies include the following:

- The positive association (correlation) between exposure and effect must be seen in individuals with definitive exposure.
- The positive association cannot be explained by bias in recording, detection, or experimental design.
- The positive association must be statistically significant.
- The positive association should show both dose and exposure duration dependence.

Information from Animal Studies

Typically, data from studies using laboratory animals must be used for some or all of the intrinsic toxicity evaluation of a chemical in humans. There are several aspects that need to be considered when interpreting the animal data, as discussed below.

Breadth and Variety of Toxic Effects The toxicological literature should be reviewed in terms of the types of effects observed in various test species. This is an important first step in chemical toxicity evaluation because:

- It identifies potential effects that might be produced in humans. To some extent, the consistency with which an effect is observed among different species provides greater confidence that this effect will occur in humans as well. An effect that occurs in some species but not others, or one sex but not the other, signals that great care will be needed in extrapolating findings in animals to humans without some form of corroborating human data.
- A comparison of effects within species (e.g., sedation vs. hepatotoxicity vs. lethality) helps establish a rank order of the toxic effects manifested as the dose increases. This

aids in identifying the most sensitive effect. Often, this effect becomes the focus of a risk assessment, since protecting against the most sensitive effect will protect against all effects. Also, comparisons of dose–response relationships within species can provide an estimation of the likelihood that one toxic effect will be seen given the appearance of another.

Mechanism of Toxicity Understanding the mechanism of action of a particular chemical helps establish the right animal species to use in assessing risk and to determine whether the toxicity is likely to be caused in humans. For example, certain halogenated compounds are mutagenic and/or carcinogenic in some test species but not others. Differences in carcinogenicity appear to be related to differences in metabolism of these chemicals because metabolism is an integral part of their mechanism of carcinogenesis. For these chemicals, then, a key issue in selecting animal data for extrapolation to humans is the extent to which metabolism in the animal model resembles that in humans. A second example is renal carcinogenicity from certain chemicals and mixtures, including gasoline. Gasoline produces renal tumors in male rats, but not female rats or mice of either sex. The peculiar susceptibility of male rats to renal carcinogenicity of gasoline can be explained by its mechanism of carcinogenesis. Metabolites of gasoline constituents combine with a specific protein, α-2μ-globulin, to produce recurring injury in the proximal tubules of the kidney. This recurring injury leads to renal tumors. Female rats and mice do not accumulate this protein in the kidney, explaining why they do not develop renal tumors from gasoline exposure. Humans also do not accumulate the protein in the kidney, making the male rat a poor predictor of human carcinogenic response in this situation.

In a sense, choosing the best animal model for extrapolation is always a catch-22 situation. Selection of the best model requires knowledge of how the chemical behaves in both animals and humans, including its mechanism of toxicity. In the situations in which an animal model is most needed (when we have little data in humans), we are in the worst position to select a valid model. The choice of an appropriate animal model becomes much clearer when we have a very good understanding of the toxicity in humans and animals, but in this situation, there is, of course, much less need for an animal model.

In addition to helping identify the best species for extrapolation, knowledge of the mechanism of toxicity can assist in defining the conditions required to produce toxicity. This is an important aspect of understanding the hazard posed by a chemical. For example, acetaminophen, an analgesic drug used in many over-the-counter pain relief medications, can produce fatal liver injury in both animals and humans. By determining that the mechanism of toxicity involves the production of a toxic metabolite during the metabolism of high doses, it is possible to predict and establish its safe use in humans, determine the consequences of various doses, and develop and provide antidotal therapy.

Dosages Tested Typically, animal studies utilize relatively high doses of chemicals so that unequivocal observations of effect can be obtained. These doses are usually much greater than those received by humans, except under unusual circumstances such as accidental or intentional poisonings. Thus, while animal studies might suggest the possibility of a particular effect in humans, that effect may be unlikely or impossible at lower dosages associated with actual human exposures. The qualitative information provided by animal studies must be viewed in the context of dose–response relationships. Simply indicating that an effect might occur is not enough; the animal data should indicate at what dosage the effect occurs and, equally importantly, at what dosage the effect does *not* occur.

Validity of Information in the Literature Any assessment of the intrinsic toxicity of a chemical begins with a comprehensive search of the scientific literature for relevant studies. While all of the studies in the literature share the goal of providing new information, the reality of the situation is that all are not equally valuable. Studies may be limited by virtue of their size, experimental design, methods employed, or the interpretations of results by the authors. These limitations are sometimes not readily apparent, requiring that each study be evaluated carefully and critically. The following are some guidelines to consider when evaluating studies:

- Has the test used an unusual, new, or unproven procedure?
- Does the test measure toxicity directly, or is it a measure of a response purported to indicate an eventual change (a pretoxic manifestation)?
- Have the experiments been performed in a scientifically valid manner?
- Are the observed effects statistically significant against an appropriate control group?
- Has the test been reproduced by other researchers?
- Is the test considered more or less reliable than other types of tests that have yielded different results?
- Is the species a relevant or reliable human surrogate, or does this test conflict with other test data in species phylogenetically closer to humans?
- Are the conclusions drawn from the experiment justified by the data, and are they consistent with the current scientific understanding of the test or area of toxicology?
- Is the outcome of the reported experiment dependent on the test conditions, or is it influenced by competing toxicities?
- Does the study indicate causality or merely suggest a correlation that could be due to chance?

Other Considerations Numerous confounders can affect the validity of information derived from animal studies and its application or relevance to human exposure to the same chemical. Issues regarding selection of the appropriate species for extrapolation are discussed in Section 23.2. Even if the selection of species is sound, certain other characteristics of the experimental animals can influence toxic responses and therefore the extrapolation of these responses to humans. Examples include the age of the animal (e.g., whether studies in adult animals are an appropriate basis for extrapolation to human children), the sex of the animal (obviously, studies limited to just male or female animals cannot address all of the potential toxicities for both sexes of humans), disease status (e.g., whether results obtained in healthy animals are relevant to humans with preexisting disease, and vice versa), nutritional status (e.g., whether studies in fasted animals accurately reflect what occurs in fed humans), and environmental conditions.

Other confounders go beyond the animal models themselves and pertain to the type of study conducted. For example, studies involving acute exposure to a chemical are usually of limited value in understanding the consequences of chronic exposure, and chronic studies generally offer little insight into consequences of acute exposure. This is because chronic toxicities are often produced by mechanisms different from those associated with acute toxicities. For this reason, good characterization of the intrinsic toxicity of a chemical requires information from treatments of varying duration, ranging from a single dose to exposure for a substantial portion of the animal's lifetime.

Information from *In Vitro* and In Silico Studies

In vitro and in silico studies are useful for predicting whether toxicity might occur as a result of exposure. The high cost of animal toxicity testing and large number of chemicals yet to be tested make these methods valuable tools for predicting hazard. *In vitro* studies include cells in culture, isolated tissues, tissue extracts or homogenates, subcellular fractions, and purified biochemical reagents (e.g., enzymes, other proteins, nucleic acids). The basis of their use in hazard identification is for determining the mechanism or mode of action and understanding how the chemical causes effects at the cellular, biochemical, and molecular level. Due to the complexity of an intact biological system, *in vitro* results cannot be extrapolated to a toxic endpoint. However, toxic effects can be predicted from these studies and verified in animal models. In silico studies include structure–activity relationships (SAR). They utilize computer modeling to predict biological activity and potency from the chemical structure. The most frequent use of in silico modeling is to predict a common mode of action for an entire class of chemicals. The Ah receptor binding ability of dioxin-like compounds was predicted in silico based on SAR.

23.3 DOSE–RESPONSE ASSESSMENT

In this portion of the risk assessment, the dose–response relationships for the toxicities of concern must be measured, modeled, or assumed, in order to predict responses to doses estimated in the exposure assessment. While dose–response relationships could theoretically be obtained for a variety of effects from each chemical of potential concern, in practice, attention is usually centered on the most sensitive effect of the chemical.

In risk assessment, two fundamentally different types of dose–response relationships are thought to exist. One is the threshold model, in which all doses below some threshold produce no effect, while doses above the threshold produce effects that increase in incidence or severity as a function of dose. The second model has no threshold—any finite, nonzero dose is thought to possess some potential for producing an adverse effect. The derivation of these two types of dose–response relationships and their use to provide estimates of risk are very different, as described in the following sections.

Threshold Models

For all toxicities other than cancer, there is some dose below which no observable or statistically measurable response exists. This dose, called the *threshold dose*, was graphically depicted in Chapter 1 (see also Figure 23.2). Conceptually, a threshold makes sense for most toxic effects. The body possesses a variety of detoxification and cell defense and repair mechanisms, and below some dose (i.e., the threshold dose), the magnitude of effect of the chemical is so small that these detoxification and defense/repair mechanisms render it undetectable.

In the most common form of threshold dose–response modeling, the threshold dose becomes the basis for establishing a "safe human dose" (SHD). Because we rarely, if ever, are able to define the true threshold point on the dose–response curve, the threshold dose is usually approximated. There are two methods for estimating the threshold. These methods include the no observable adverse effect level (NOAEL)/lowest observable adverse effect level (LOAEL) approach and the benchmark dose (BMD) approach. The preferred method is the BMD approach, which derives a threshold dose based on the dose–response curve. If data are not amenable to the BMD approach, then the NOAEL/LOAEL approach is utilized.

In the NOAEL/LOAEL approach, the more desirable method uses the highest reported dose or exposure level for which no toxicity was observed. This dose, known as the "NOAEL," is considered for practical purposes to represent the threshold dose. This prevents underestimating the toxicity of a chemical. Sometimes, the available data do not include a NOAEL; that is, all of the doses tested produced some measurable toxic effect. In this situation, the lowest dose producing an adverse effect, termed the

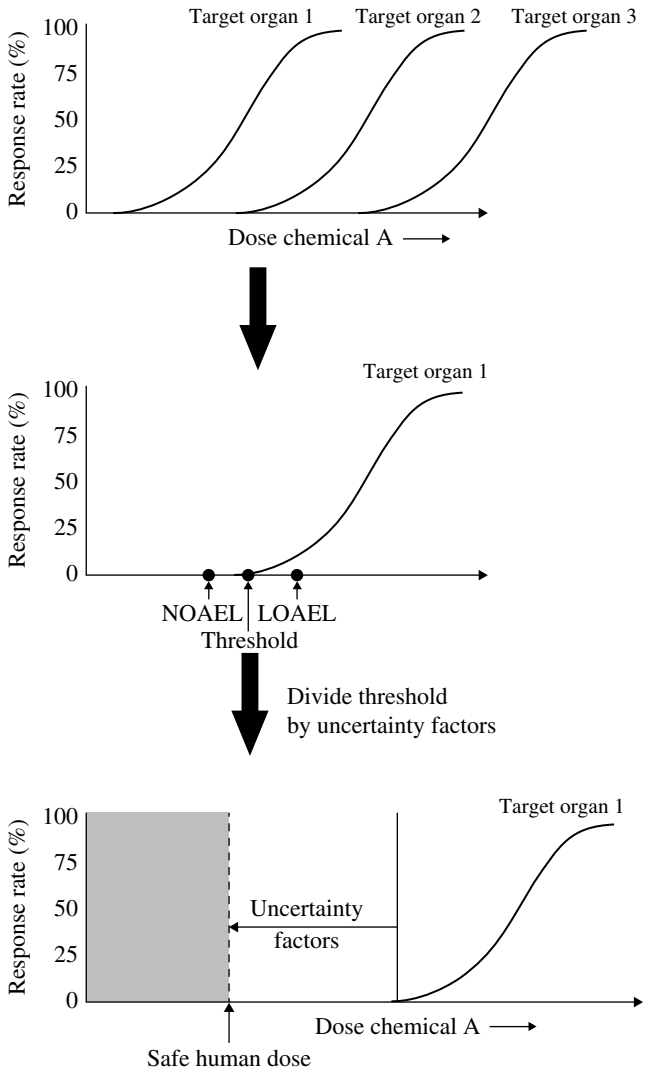


FIGURE 23.2 Estimation of a safe human dose (SHD). The first step is identification of the target organ or effect most responsive to the chemical (in this case, target organ 1 in the upper panel). Dose–response data for this effect are used to identify no observable adverse effect level (NOAEL) and/or lowest observable adverse effect level (LOAEL) doses in order to approximate the threshold dose. Either the NOAEL or LOAEL is divided by a series of uncertainty factors to generate the SHD.

"LOAEL," is identified from the dose–response data. The threshold dose will lie below, and hopefully near, this dose. A threshold dose is then projected from the LOAEL, usually by dividing the LOAEL by a factor of 10 (see Figure 23.2). There are several limitations to the NOAEL/LOAEL methodology. One limitation is that the ability of the NOAEL to approximate the threshold dose is dependent on dose selection and spacing in available studies, and in many cases, these are not well suited to determining the threshold. If the doses are spaced far apart, the NOAEL may be much lower than the actual threshold dose and result in an overestimation of toxicity. A second limitation is that the approach fails to consider the shape or slope of the dose–response curve, focusing instead on results from one or two low doses exclusively. This is especially important at lower, environmentally relevant concentrations where small changes in dose can result in large changes in effect. Another limitation is that studies with small numbers of animals may result in higher thresholds since there is not enough power in those studies to identify effects that occur with low frequencies. Therefore, poor study design can be rewarded with a higher threshold dose.

The second method for estimating a threshold dose is the BMD approach. This method utilizes all of the dose–response data and is not dependent on any single data point. In this approach, dose–response data for the toxic effect of concern are fit to a mathematical model, and the model is used to determine the dose corresponding to a predetermined benchmark response. For most quantal data, the dose at which 10% of the population exhibits a response (effective $dose_{10}$ (ED_{10})) is chosen as the benchmark response. Exceptions include reproductive data (5% level) and human data (1% level) for which lower benchmarks are utilized. For continuous data, the benchmark response is a 10% change in endpoint that is considered to be biologically significant or a change of the treated mean equal to one standard deviation from the control mean. As an example, dose–response data might be used to determine the dose required to produce a 10% incidence of liver toxicity from mice treated with a chemical. This dose would be referred to as the ED_{10} or dose effective in producing a 10% incidence of effect. Often, for regulatory purposes, statistical treatment of the data is used to derive upper and lower confidence limit estimates of this dose. The more conservative of these is the lower confidence limit estimate of the dose, which in this case would be designated as the $BMDL_{10}$ (see Figure 23.3). In order to develop an SHD from the ED_{10} or the $BMDL_{10}$, a series of uncertainty factors would be applied, analogous to the NOAEL approach. In a sense, the BMD approach is like extrapolating an SHD from a NOAEL, except the BMD is much more rigorously defined.

The BMD approach works best if there are response data available for a variety of doses. In order to derive an accurate estimate of the threshold dose utilizing the BMD approach, a statistically or biologically significant dose-related trend is necessary. Without a trend, the software will not be able to accurately model the data. Additionally, if there is no NOAEL, the LOAEL must be near the true threshold. Otherwise, there is too much uncertainty in the models (the BMDL will be model dependent), and the software will not be able to reproduce the shape of the curve in the threshold region with any certainty. In this case, the BMD approach would not provide any additional information as to the location of the threshold dose. In these instances, the NOAEL/LOAEL approach should be used.

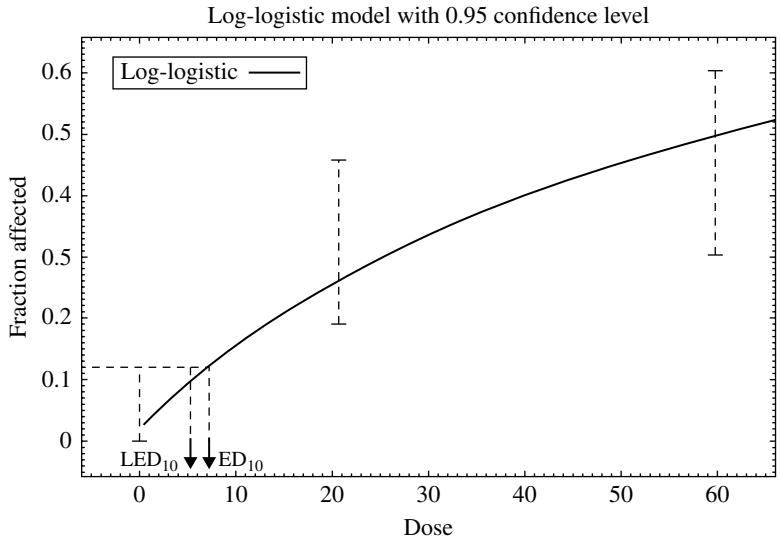


FIGURE 23.3 Derivation of the benchmark dose (BMD). Dose–response data for the toxic effect of concern are fit to a mathematical model, depicted by the solid line. This model is used to determine the effective dose (ED) corresponding to a predetermined benchmark response (BMR) (e.g., 10% of the animals responding), which is termed the BMD. Statistical treatment of the data can be used to derive the lower confidence limit estimate of the BMD, termed the BMDL. Either the BMD or BMDL may be used to represent the point of departure, although the more conservative BMDL is typically used for regulatory purposes. *Source:* Adapted from USEPA (2012).

From the estimates of the threshold dose, an SHD can be calculated. Different agencies have different terminologies that they apply to the SHD; the U.S. EPA refers to this dosage as a "reference dose" (RfD), or, if it is in the form of a concentration of chemical in air, as a "reference concentration" (RfCs). Other agencies have adopted different terminologies; for example, the U.S. FDA uses the term "allowable daily intake" (ADI). The basic concept is the same, and the approach to the development of an SHD is relatively simple, as illustrated in the flow diagram in Figure 23.2. Because a chemical may produce more than one toxic effect, the first step is to identify from the available data the adverse effect that occurs at the lowest dose. Second, the threshold dose or some surrogate measure of the threshold dose (e.g., the NOAEL or LOAEL reduced by some amount) is identified for the most sensitive toxic endpoint. The threshold dose (or its surrogate measure) is then divided by an uncertainty factor to derive the SHD, and this dose can then be converted into an acceptably safe exposure guideline for that chemical.

Calculating Safety for Threshold Toxicities: The SHD Approach

The calculation of an SHD essentially makes an extrapolation on the basis of the size differential between humans and the test species. This extrapolation is based on a dosimetric adjustment factor that accounts for toxicokinetic and some toxicodynamic differences between species. The calculation is similar to the following:

$$\text{SHD} = \frac{\text{NOAEL}\,(\text{mg/kg per day}) \times (\text{BW}_a / \text{BW}_h)^{1/4}}{\text{UF}}$$
$$= N\,\text{mg/kg}\cdot\text{day}$$

where

NOAEL = threshold dose or some other no observable adverse effect level selected from the no-effect region of the dose–response curve;

SHD = safe human dose;

UF = the total uncertainty factor, which depends on the nature and reliability of the animal data used for the extrapolation;

N = number of milligrams consumed per kilogram per day;

BW_a = body weight of the animal; and

BW_h = human body weight.

Typically, the uncertainty factor used varies from 10 to 10,000 and is dependent on the confidence placed in the animal database as well as whether there are human data to substantiate the reliability of the animal no-effect levels that have been reported. Of course, the number calculated should use chronic exposure data if chronic exposures are expected. This type of model calculates one value, the expected safe human dosage, that regulatory agencies have referred to as either the ADI or the RfD. Exposures, which produce human doses that are at or below these safe human dosages (ADIs or RfDs), are considered safe.

Example Calculation Pentachlorophenol (PCP), a general-purpose biocide, will be used as an example of how to derive

a safe human dosage. A literature review of the noncarcinogenic effects of PCP has shown that the toxicological effect of greatest concern is its hepatotoxic effects in test animals. The PCP LOAEL for these effects has been reported to be 1.5 mg/kg daily. Using the formulas shown in the previous text and an uncertainty factor of 300, an SHD could be calculated as follows:

$$\text{SHD} = \frac{1.5\text{mg/kg daily} \times (12\text{kg}/70\text{kg})^{1/4}}{300}$$

$$\text{SHD} = 0.0032\text{mg/kg} \cdot \text{day}$$

Once the SHD has been estimated, it may be necessary to convert the dose into a concentration of the chemical in a specific environmental medium (air, water, food, soil, etc.) that corresponds to a safe exposure level for that particular route of exposure. That is, while some dose (in mg/kg · day) may be the total safe daily intake for a chemical, the allowable exposure level of that chemical will differ depending on the route of exposure and the environmental medium in which it is found.

Uncertainty Factor The uncertainty factor is really a composite of several uncertainty factors intended to address weaknesses in the data or uncertainties in extrapolation from animals to humans. These uncertainties arise because of our inability to directly measure the actual human threshold dose. The weaker the data set available for evaluation (few studies, limited doses tested, etc.) and the more assumptions required, the greater the uncertainty that the NOAEL or LOAEL from the literature actually represents the threshold dose in humans. The purpose of dividing the NOAEL, LOAEL, or BMD by uncertainty factors is to ensure that the SHD used in the risk assessment is below the actual human threshold dose for toxicity for all individuals in the exposed population, thereby avoiding any underestimation of risk. The greater the uncertainty associated with the data, the larger the uncertainty factor required to insure protection.

The general rationale for selecting the size of the uncertainty factor for a particular area of uncertainty is as follows:

- UF_A—An uncertainty factor of up to 10 is applied in extrapolating toxicity data from one species to another. It is used to account for the possibility that humans are more sensitive to toxicity than the test species. A factor of 10 is utilized as the default value, and a factor of 3 is utilized if the study species is a nonhuman primate or if toxicodynamic and toxicokinetic data allow the calculation of a human equivalent dose (e.g., physiologically based pharmacokinetic (PBPK) modeling, species scaling).
- UF_H—An uncertainty factor of up to 10 is used to account for variability in sensitivity to toxicity among subjects. An uncertainty factor of 10 is applied to ensure that the final toxicity value is protective for sensitive individuals within a population. An uncertainty factor of 3 is utilized if the data is from a sensitive subpopulation known to be more susceptible to the adverse effect. An uncertainty factor of 1 is utilized if human data are available from a particularly vulnerable subpopulation.
- UF_S—An uncertainty factor of up to 10 might be applied if only subchronic data are available. It is possible under these circumstances that the threshold dose for longer exposures might be lower, and this uncertainty factor is intended to protect against this possibility. A factor of 3 is often utilized for an exposure duration that is greater than subchronic, but less than chronic (e.g., 1 year in rodents). An uncertainty factor of 1 is utilized when chronic data are available.
- UF_L—As discussed earlier, an uncertainty factor of up to 10 may be applied if the only value with which to estimate the threshold dose is a LOAEL value. Division by this uncertainty factor is meant to accomplish a reduction in the LOAEL to a level at or below the threshold dose. An uncertainty of 1 is applied when the BMD is utilized as the threshold dose.
- UF_D—An additional uncertainty factor of up to 10 is applied if the overall quality of the database is poor, the number of animal species tested is few, the number of toxic endpoints evaluated is small, or the available studies are found to be deficient in quality. An uncertainty factor of 3 is utilized if either a prenatal toxicity or two-generation reproduction study is absent from the database. If both are absent, an uncertainty factor of 10 should be applied.
- MF—The development of some SHDs incorporates a modifying factor to account for deficiencies in the data set not covered by the other uncertainty factors. In 2002, the U.S. EPA discontinued the use of modifying factors, stating they are sufficiently incorporated in the general database uncertainty factors.

These uncertainty factors are multiplicative; that is, an uncertainty factor of 10 for sensitive individuals combined with an uncertainty factor of 10 for extrapolation of data from animals to humans results in a total uncertainty factor of 100 (10×10). Total uncertainty factors applied to develop an SHD commonly range between 300 and 1000, and values up to 10,000 or more are possible, although regulatory agencies may place a cap on the size of compounded uncertainty factors (e.g., a limit of 3000).

In the example calculation for PCP earlier, an uncertainty factor of 300 was utilized to calculate the SHD. Uncertainty factors were used to account for animal to human extrapolation, variability among humans in sensitivity, and the use of a LOAEL (UF_A of $10 \times UF_H$ of $10 \times UF_S$ of $1 \times UF_L$ of $3 \times UF_D$ of $1 = 300$).

Quantifying Noncancer Risk

Although the term "risk" often implies probability of an adverse event, the threshold approach to assessing chemical risk does not result in risk expression in probability terms. This approach is instead directed to deriving a safe limit for exposure and then determining whether the measured or anticipated exposure exceeds this limit. All doses or exposures below this "safe level" should carry the same chance that toxicity will occur—namely, zero. With this model, the acceptability of the exposure is basically judged in a "yes/no" manner. The most common quantitative means of expressing hazard for noncancer health effects is through a hazard quotient (HQ). Agencies such as the U.S. EPA calculate an HQ as the estimated dose from exposure divided by their form of the SHD, the RfD:

$$\mathrm{HQ} = \frac{D}{\mathrm{RfD}} \left(\text{or } \mathrm{HQ} = \frac{D}{\mathrm{SHD}} \right)$$

where

HQ = hazard quotient,

D = dosage (mg/kg · day) estimated to result from exposure via the relevant route, and

RfD = reference dose (mg/kg · day).

Interpretation of the HQ is relatively straightforward if the value is less than one. This means that the estimated exposure is less than the SHD and no adverse effects would be expected under these circumstances. Interpretation of HQ values greater than one is more complicated. A value greater than one indicates that the estimated exposure exceeds the SHD, but recall that the SHD includes a number of uncertainty factors that impart a substantial margin of safety. Therefore, exposures that exceed the SHD, but lie well within this margin of safety, may warrant further analysis but are unlikely to produce adverse health effects.

Dose–response relationships can vary from one route of exposure to another (e.g., a safe dose for inhalation of a chemical may be different from a safe dose for its ingestion). As a result, a given chemical may have different SHDs for different routes of exposure. Since individuals are often exposed to a chemical by more than one route, separate route-specific HQ values are calculated. For example, the estimated inhalation dose would be divided by the SHD for inhalation to calculate an HQ for inhalation, while the estimated dose received from dermal contact would be divided by a dermal SHD to derive the HQ for this route of exposure. Typically, the HQ values for each relevant route of exposure are summed to derive a hazard index (HI) for that chemical. Interpretation of the HI is analogous to the HQ—values less than one indicate that the safe dose has not been exceeded (in this case, by the aggregate from all exposure routes). A value greater than one suggests that effects are possible, although not necessarily likely. The HI is also a means by which effects of different chemicals with similar toxicities can be combined to provide an estimate of total risk to the individual.

Another means to convey the relationship between estimated and safe levels of exposure is through calculation of a margin of exposure. This is most often used in the context of the BMD approach. The margin of exposure is the BMD divided by the estimated dose. An acceptable margin of exposure is usually defined by the uncertainty factors applied to the BMD. If, for example, available data suggest that a total uncertainty factor of 1000 should be applied to the BMD for a specific chemical and effect, and the margin of exposure for that chemical is greater than 1000 (i.e., the estimated dose is less than the BMD divided by 1000), the exposure would be regarded as safe.

A Special Case: Assessing Risk from Lead Exposure

The aforementioned methods are almost universally applied in assessing the potential for noncancer health effects. There is, however, one exception for which a radically different approach is used: the evaluation of noncancer effects from lead in children. In 2012, the Centers for Disease Control and Prevention (CDC) recommended blood lead concentrations in children should not exceed 5 μg/dl in order to avoid intellectual impairment (this is a decrease from their previous recommendation of 10 μg/dl). Thus, the main objective in lead risk assessment is to determine whether childhood lead exposure is sufficient to result in a blood lead level that causes adverse effects.

To predict blood lead levels from environmental exposure, the U.S. EPA has developed a PBPK model known as the "integrated exposure uptake biokinetic model for lead in children" (IEUBK). The IEUBK model has four basic components (i.e., exposure, uptake, biokinetics, and probability distribution) and uses complex mathematics to describe age-dependent anatomical and physiological functions that influence lead kinetics. The model predicts the blood concentration (the dose metric most closely related to the health effect of interest) that results from an endless variety of exposure scenarios that can be constructed by the risk assessor (i.e., exposure to various concentrations of lead in soil, dust, water, food, and/or ambient air). The model also predicts the probability that children exposed to lead in environmental media will have a blood lead concentration exceeding a health-based level of concern (see Figure 23.4). It is important to note that the U.S. EPA has not yet decreased their health-based level of concern and currently utilizes 10 μg/dl as specified in the 1994 *Revised Interim Soil Lead Guidance for CERCLA Sites and RCRA Corrective Action Facilities.* However, the IEUBK model allows the user to choose the blood lead level of concern. The IEUBK approach is rather unique because it is among the few approaches that

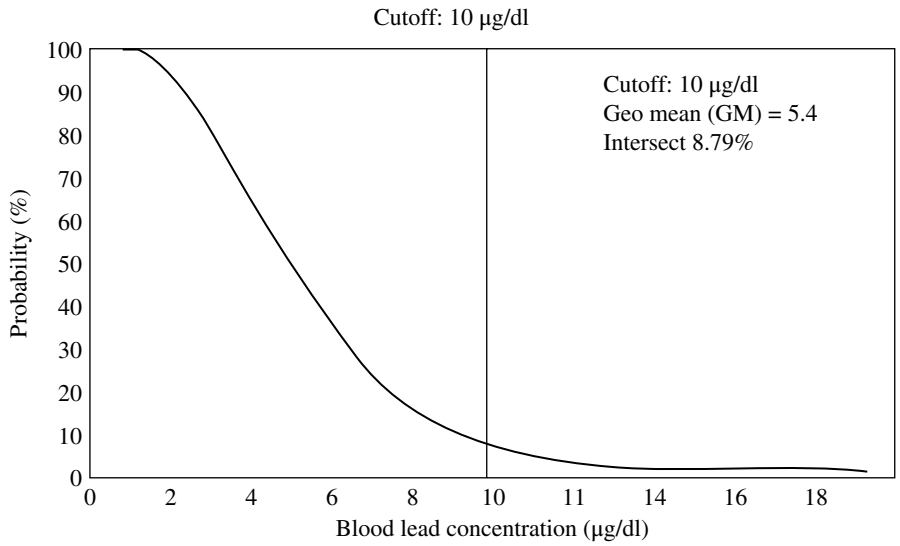


FIGURE 23.4 Example of output from the IEUBK model. The curve displays the cumulative probability of developing a blood lead concentration at varying media concentrations as a result of the specified exposure. In this example, there is a probability of virtually 100% that the modeled exposure will result in a blood lead concentration greater than 1 µg/dl, but only about a 9% probability that the blood lead concentration will exceed 10 µg/dl.

rely on an internal dose metric (i.e., blood lead level) and PBPK modeling for risk assessment purposes.

Nonthreshold Models for Assessing Cancer Risks

Conceptual Issues The nonthreshold dose–response model is typically reserved for cancer risk assessment. The assumption by regulatory agencies that chemical carcinogenesis has no dose threshold began several decades ago. This assumption was initially based largely on empirical evidence that radiation-induced cancer had no threshold and on the theory that some finite amount of DNA damage was induced by all doses of radiation. Smaller doses simply carried smaller risks, but all doses were assumed to carry some mathematical chance of inducing cancer. Following this lead, theories of chemical-induced carcinogenesis began to evolve along the same lines, centering on effects of highly reactive, DNA damaging carcinogens. It was presumed that, like radiation, chemical carcinogens induced cancer via mutations or genetic damage and therefore had no thresholds. So, like radiation before it, chemical-induced carcinogenesis was assumed to carry some quantifiable risk of cancer at any dose.

If viewed somewhat simplistically, a biologic basis for the absence of a practical threshold for carcinogens can be hypothesized. If one ignores the DNA repair processes of cells, or assumes that these protective processes become saturated or overwhelmed by "background" mutational events, it can be postulated that some unrepaired genetic damage occurs with each and every exposure to a carcinogenic substance. As this genetic damage is presumed to be permanent and carry the potential to alter the phenotypic expression of the call, any amount of damage, no matter how small, might carry with it some chance that the affected cell will ultimately evolve to become cancerous.

With this viewpoint, scientists and regulatory agencies initially proposed that the extrapolation of a cancer hazard must be fundamentally different from that used to extrapolate noncancer hazards, and cancer risk assessment models become probability based. In contrast to assessing the risk of noncancer health effects, where the dose at which no toxic effect will occur is determined, cancer risk assessment is a matter of assigning probabilities of cancer to different doses. The determination of safety or a safe dose is then a matter of deciding what cancer risks are so small that they can be regarded as de minimis or inconsequential.

Determining the relationship between carcinogen dose and cancer risk is very difficult for a number of reasons. One reason is that the concept of latency complicates the interpretation of dose–response relationships for carcinogens. *Latency* is the interval of time between the critical exposure and the ultimate development of disease. While noncancer effects tend to develop almost immediately or very soon after a toxic dose is received, cancer may not develop until an interval of 20 years or more has elapsed. For some carcinogens, increasing the dose shortens the latency period, causing tumors to develop more quickly. A positive carcinogenic response can then be thought of in two ways: as increased numbers of tumors or subjects with tumors or as a decrease in the time to appearance of tumors. The latter is important, because a dose capable of producing tumors has no consequence if the time required for the tumors to develop exceeds the remaining lifespan of a human or animal.

Another problem is that the critical portion of the dose–response curve for most risk assessments, the low-dose region applicable to most environmental and occupational exposures, is one for which empirical data are not available. Chronic cancer bioassays in animals are expensive and seldom test more than two or three doses. Also, cost limits the number of animals tested to about 50 or less per dose group. With this group size, only tumor responses of about 10% or more can be detected with statistical significance. Detection of the kinds of cancer responses that might be of interest to the risk assessor, for example, a response of 0.1, 0.001 or 0.00001% (10^{-3}, 10^{-5}, or 10^{-6}, respectively), is therefore beyond the capabilities of these experiments. Consequently, the doses needed to produce these cancer responses are not determined. Expanding the number of animals routinely tested is not economically feasible, and even very large studies may not eliminate this problem. One attempt to test the utility of using larger dose groups, the so-called "megamouse" experiment, was still unable to increase the sensitivity of measurement beyond about 1%, even though almost 25,000 animals were used in this experiment. In short, animal cancer bioassays will typically provide only one or two dose–response points, and these points are always several orders of magnitude above the range of small risks/doses in which we are ultimately interested.

Because low-dose responses cannot be measured, they must be modeled. There are three types of models:

1. The first category of models consists of the "mechanistic" models. These are dose–response models that attempt to base risk on a general theory of the biological steps that might be involved in the development of carcinogenesis. Examples of mechanistic models include the early "one-hit" and the subsequent "multihit" models for carcinogenesis. These models were based on assumptions concerning the number of "hits" or events of significant genetic damage that were necessary to induce cancer. A related model, the "linearized multistage" (LMS) model of carcinogenesis, is based on the theory that cancer cells develop through a series of different stages, evolving from normal cells to cancer cells that then multiply.
2. The second category of cancer extrapolation models includes the "threshold distribution" models. Rather than attempting to mimic a particular theory of carcinogenesis, these models are based upon the assumption that different individuals within a population of exposed persons will have different risk tolerances. This variation in tolerance in the exposed population is described with different probability distribution of the risk per unit of dose. Models that fall within this category include the probit, the logit, and the Weibull.
3. The third category of model is the "time-to-tumor" model. This type of model bases the risk or probability of getting cancer on the relationship between dose and latency. With this model, the risk of cancer is expressed temporally (in units of time), and a safe dose is selected as one where the interval between exposure and cancer is so long that the risk of other diseases becomes of greater concern.

Each of these models can accommodate the assumption that any finite dose poses a risk of cancer, the essential tenet of a nonthreshold model. However, the shape of the dose–response curve in the low-dose region can vary substantially among models (see Figure 23.5). Because the shape of the dose–response curve in the low-dose region cannot be verified by measurement, there is no means to determine which shape is correct. A simple example of the impact of choosing one cancer extrapolation model over another is given in Table 23.1, which compares the results of dose–response modeling using three different models where it was assumed in each model that a relative dose of 1.0 produced a 50% cancer incidence. The results generated by all three models are essentially indistinguishable at high doses where the animal cancer incidence might be observable, and so one would conclude that they all "fit" the experimental data equally well. However, when modeling the risks associated with lower doses, the dose/risk range in which regulatory agencies and risk assessors are most frequently interested,

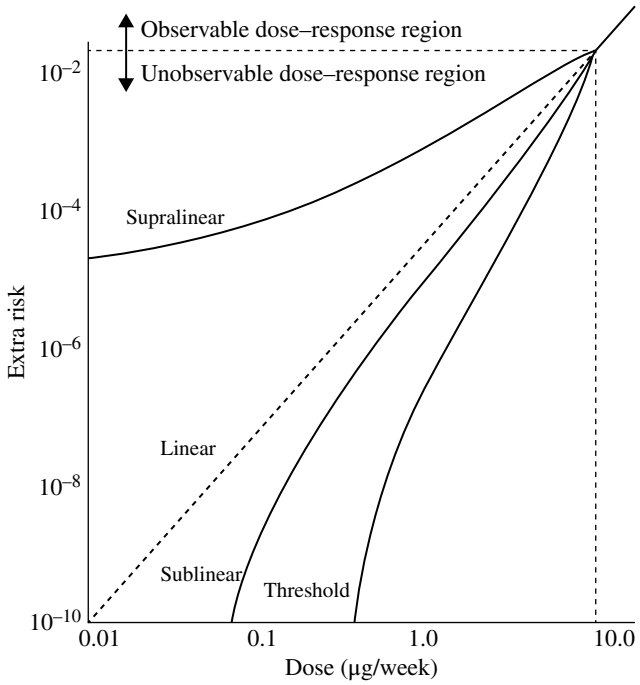


FIGURE 23.5 Four different extrapolation models applied to the same experimental data. All fit the data equally well in the observable range, but each yields substantially different risk estimates in the low-dose region most applicable to occupational and environmental exposures. *Source:* Adapted from NRC (1983).

TABLE 23.1 Expected Risk (Cancer Incidence) Calculated by Three Models when a Relative Dose of 1.0 Is Assumed to Cause a 50% Tumor Incidence in Test Animals

	Modeled Tumor Incidence (%)		
Relative Dose	Lognormal	Log-Logistic	One-Hit
16	98	96	100
4	84	84	94
1[a]	**50**[a]	**50**[a]	**50**[a]
1/4	16	16	16
1/16	2	4	4
1/100	0.05	0.4	0.7
1/1000	0.00035	0.026	0.07
1/10,000	0.0000001	0.0016	0.007

Source: Adapted from Office of Technology Assessment (1981).
[a] Boldface numbers represent the only data assumed for each model; all other tumor incidences were calculated from this single dose–response point.

TABLE 23.2 The Estimated Impact of Six Conservative U.S. EPA Assumptions on Agency Risk Assessments

Factor	Range of Possible Overstatement of Estimated Cancer Risk[a]
Body weight versus surface area as a scalar for an interspecies extrapolation	2–12
MLE[b] versus 95% UCL[c] for the cancer slope factor	1–3
Malignant tumors only versus malignant + benign	1–2
Average species sensitivity versus most sensitive species	2–5
Pharmacodynamics versus effective dose	1–6
Risks at shorter than equilibrium buildup time	2–5
Total risk exaggeration	15–10,800

Source: Adapted from Barnard (1994) and based on information supplied by Dr. E. Anderson.
[a] Instead of presenting ranges of possible overstatement in cancer risk, the Barnard paper presents ranges of possible reduction in estimated cancer risk if the alternative factors to the current default factors are applied.
[b] Maximum likelihood exposure.
[c] Upper confidence limit.

there is a wide divergence in the risk projected by each model for a given low dose. In fact, at 1/10,000th of the dose causing a 50% cancer incidence in animals, the risks predicted by these three models produce a 70,000-fold variation in the predicted response.

Regulatory agencies utilize cancer risk estimates in regulating carcinogens, but they are faced with many models that yield a wide range of risk estimates. In the absence of any scientific basis to determine which is most correct, they must make a science policy decision in selecting the model to use. Generally, in the face of this uncertainty, they have selected models that tend to provide higher estimates of risk particularly when combined with conservative exposure assumptions (see Table 23.2). This is consistent with their mission to protect public health, and consequently the need to avoid underestimating risks. For example, the U.S. EPA has historically used conservative models such as the one-hit of LMS model in calculating cancer risks from exposure to all carcinogens. These models assume linearity in the low-dose range and, as shown in Table 23.1, tend to require a larger reduction in dose to attain a certain low level of risk relative to other models.

Extensive research in the area of chemical carcinogenesis indicates that many chemical carcinogens act via epigenetic or promotional mechanisms that, like noncancer toxicities, do not involve or require genetic damage. It has been proposed that these mechanisms and carcinogenic responses should have thresholds. Similarly, numerous enzyme systems have been identified as responsible for maintaining the integrity of the genetic code. These repair enzymes and pathways could provide an effective dose threshold for even those carcinogens whose mechanism is believed to involve some mutational event or other form of genetic damage.

In recent years, debates involving the actual shape of the dose–response curve for carcinogens in the low-dose region, and the issue of thresholds for carcinogens, have caused scientists and regulatory agencies to reevaluate earlier cancer risk assessment methodologies. Out of this reevaluation has come a movement to adopt two major policy changes in the cancer risk assessment methodologies employed by regulatory agencies. One proposed change is to use risk extrapolation models that make fewer assumptions about the shape of the dose–response curve (e.g., the BMD and margin-of-exposure method). Data within the observation range can be used to develop a "point of departure," the critical point for extrapolating responses in the low-dose range. For dose–response relationships assumed to have no threshold, the simplest extrapolation model is used: a straight line drawn between the point of departure and zero. The second proposed change is to allow for the consideration and use of nonlinear and threshold models for carcinogens where empirical and mechanistic evidence argues strongly that this type of dose–response model is appropriate for a particular chemical. In this situation, risk of cancer would be evaluated in a manner analogous to noncancer health effects, such as through calculation of a margin of exposure.

Generating Cancer Risk Estimates Estimating the lifetime cancer risk associated with a particular dose is a relatively simple mathematical process. Because most regulatory agencies such as the U.S. EPA use the conservative assumption that cancer risk should be modeled via a linear, nonthreshold

model, the risk associated with a particular dose is calculated by the following formula:

$$R = D \times \text{CSF} \quad \text{or} \quad R = \text{LADD} \times \text{CSF}$$

where

R = risk,

D = dose (normally expressed as the lifetime average daily dose (LADD) (mg/kg · day)), and

CSF = cancer slope factor (the slope of the dose–response curve in units of $(\text{mg/kg} \cdot \text{day})^{-1}$).

With this equation, the total dose the individual or population has accumulated during their entire exposure interval is first converted into a LADD, a dose that if received every day for a lifetime would be equivalent to the total dose accumulated during the actual exposure period. For example, if the exposure assessment projected a daily dosage of 70 mg/kg · day for a 30-year exposure interval, then the LADD assuming a 70-year lifespan would be 30 mg/kg · day (i.e., 70 mg/kg · day × 30 years ÷ 70 years = 30 mg/kg · day). The dose is expressed in units of mg/kg · day, and the CSF is in units of reciprocal mg/kg · day or $(\text{mg/kg} \cdot \text{day})^{-1}$. If the chemical in this example has a CSF of 0.001 $(\text{mg/kg} \cdot \text{day})^{-1}$ (in scientific notation a value of 1.0×10^{-3} $(\text{mg/kg} \cdot \text{day})^{-1}$) and the LADD derived during the exposure assessment was 0.03 mg/kg · day (3.0×10^{-2} mg/kg · day), the risk would be as follows:

$$R = 0.001\,\text{mg/kg day} \times 0.03\,(\text{mg/kg day})^{-1} = 0.00003$$

Which can also be written as

$$R = \frac{3}{100{,}000} \quad \text{or} \quad R = 3.0 \times 10^{-5}$$

In this example, the risk estimate represents a 3/100,000 chance or mathematical probability that a cancer will develop from exposure. It should also be noted, however, that because regulatory agencies strive for conservative, health-protective risk calculations, the CSF used is statistically an upper-bound estimate of the dose–cancer relationship. The true cancer risk of the chemical at this dose may be much less than that calculated and, in fact, could be as low as zero.

Dose Metrics

A common issue for both threshold and nonthreshold dose–response relationships is the metric used to express dose. The dose metric is important because animal data must often be used as a surrogate for dose–response information in humans. Humans are, of course, much different in size than most laboratory animals. How then should doses be scaled between one animal species and another, and between animals and humans?

One can improve the accuracy of SHD calculations by starting with the best "dose metric" (measure of the dose) for the actual amount of chemical required to induce toxicity in the most sensitive target organ. Most dose information in animal studies is reported in terms of the "applied dose" (the dose administered to the whole animal). Remember, however, that it is only the "absorbed dose" (the amount of the chemical actually absorbed into the body) that is eligible for inducing toxicity. Further, from the dose that is absorbed, it is the dose that reaches the target tissue that is most important in determining the extent of response. The relationship between applied dose and target organ dose can be different among species, due to differences in metabolism and/or distribution of the chemical within the body, leading to important differences in apparent dose–response relationships (i.e., those based strictly on applied dose). One approach used to enhance extrapolation among species is PBPK modeling. Using PBPK models, scientists are able to predict target organ doses of a chemical (or a critical metabolite, if that is important for toxicity) in test species and humans. With this information, corrections can be made for pharmacokinetic differences among species, leading to better extrapolation of dose–response relationships. The principal limitation of PBPK analyses is that they are data intensive, and PBPK models have been constructed and validated for only a few chemicals (see Chapter 3).

Since PBPK models are not often available, simpler approaches to extrapolating doses must be used in most situations. One of the simplest approaches is to convey doses per unit body weight. Larger animals (or humans) are assumed to require larger doses to produce the same toxic effect in proportion to their body weight. This is the dose metric most commonly used when extrapolating information on health effects among species. In biology, empirical observations suggest that many biochemical and physiological processes seem to scale among species according to surface area, while differences in others seem to correspond more closely to changes in weight. The correct scaling for doses is not entirely obvious and could conceivably be different for different chemical classes or toxicological effects. The current recommendation from the U.S. EPA is that scaling for both noncarcinogen and carcinogen doses uses a factor intermediate between body weight (or body weight raised to the power of 1) and surface area (body weight raised to the power of 0.67); that is, body weight raised to the power of 0.75.

Does the choice of scaling factor really make a difference? To illustrate the answer, consider the extrapolation of dose information between a mouse and a human. If a dose for a noncancer effect in a mouse were converted to a human dose based on surface area, rather than on body weight, the SHD dose would be reduced by a factor of 12 to 14. On the other hand, switching from surface area scaling to body weight

scaling for carcinogenicity data would result in a 12- to 14-fold decrease in cancer risks estimated from the same dose–response information. The difference in use of body weight versus surface area for extrapolating between rats and humans is not as large (about six-fold), but still might be considered significant.

23.4 EXPOSURE ASSESSMENT: EXPOSURE PATHWAYS AND RESULTING DOSAGES

Exposure assessment can be defined as the measurement or estimation of the amount or concentration of a chemical(s) coming into contact with the body at potential sites of entry (e.g., skin, lung, GI tract). Not only are the amount and route of exposure concerns, but so too are the exposure duration, exposure frequency, and any factors that modify the ability of the chemical to traverse the portals of entry into the body. In cases where a potential chemical hazard exists, exposure assessment is an obligatory part of the risk assessment process. Without exposure, even the most hazardous chemical poses no risk. Conversely, excessive exposure to minimally hazardous chemicals may pose an unacceptable risk. Therefore, risk assessment requires that toxicity and exposure assessments be coupled. Initially, exposure assessments should identify all potential exposure pathways and assess their completeness, after which the quantification of exposure via each relevant pathway should be determined.

Exposure pathways consist of four basic parts: (1) source of contamination, (2) contaminated media, (3) contact with the contaminated media, and (4) a route of exposure. All of these parts must be present to produce an exposure. If one of these parts is absent, there is no exposure and no risk. Soil, sediment, groundwater, surface water, air, dust, and biota may all function as the contaminated media. They become contaminated directly by a chemical release or indirectly by contact with another contaminated media. Exposure pathways are usually illustrated in the conceptual site model (CSM). The CSM is a visual representation of the complete and incomplete exposure pathways at a release. Depending on the complexity of the CSM, exposure routes may also be represented. The main exposure routes include ingestion, dermal absorption, and inhalation (both vapors and particulates).

There are two basic methods for quantifying exposure: exposure measurement and exposure modeling. Measurement results in the most accurate and realistic exposure data, but fully characterizing variable exposures that might occur to multiple receptors via multiple pathways for an extended period of time is seldom feasible. In general, the measurement of occupational exposure is easier than environmental exposure, since the former usually occurs in a confined facility, whereas the latter involves more complex time–activity patterns. Also, occupational exposure limits are typically based on 8 h time-weighted averages and 15 min short-term exposure limits, making monitoring for compliance purposes manageable from a time standpoint. Examples of personal exposure measures include analyzing a person's intake of food and water and the contaminants therein, collecting and analyzing a urine sample at the end of a work shift, and measuring airborne exposure with a portable sampling device suspended in a person's breathing zone. Where environmental exposure to a large population is at issue, personal exposure monitoring is not a realistic approach. Rather, environmental media suspected of being contaminated are sampled and population-based assumptions about intake rates are made. Personal questionnaires and time–activity logs are helpful in making accurate exposure estimates within a large population.

In those cases where monitoring data are unavailable or inadequate for exposure assessment, models are used to simulate the behavior of chemicals and predict their concentrations in the environment. Hundreds of such exposure models exist, including atmospheric models, surface water models, groundwater models, and food chain models. All of these models are limited by uncertainty in the data input, as well as uncertainty to the predictive capability of a generic model for a specific exposure scenario. In recognition of this uncertainty, models used for regulatory purposes tend to provide liberal estimates of exposure that may overstate risk. Whenever models are used, an attempt should always be made to collect site- or situation-specific data for the purpose of model validation. Despite their limitations, exposure models are of value in that they can make predictions for an unlimited number of exposure scenarios and predict past and future exposures. Exposure measurement, on the other hand, is limited to the present.

Exposure or concentration is often expressed in units of $\mu g/m^3$ (air), μg/l (water), or $\mu g/cm^2$ (skin). Air and water concentrations are also frequently reported in parts per million (ppm) or parts per billion (ppb) units that reflect the weight or volume of chemical per unit volume of the carrier medium. For some chemicals, risk can be directly calculated from these concentration terms using unit risk factors that are expressed as risk per μg/l (water) or risk per $\mu g/m^3$ (air). In such cases, risk is simply the product of the chemical concentration and the unit risk factor. Exposure or concentration data can also be directly compared to many occupational (e.g., OSHA permissible exposure limits (PELs) and the American Conference of Governmental Industrial Hygienists (ACGIH) Threshold Limit Values (TLVs®)) and environmental (U.S. EPA Maximum Contaminant Levels [MCLs], National Ambient Air-Quality Standards [NAAQS], and RfCs) exposure standards that have risk considerations inherent in their derivation.

While some exposure to a hazardous chemical is required in order to have risk, it is dose that relates more closely to the toxic response. Dose, often expressed in units of mg/kg body weight · day, is the amount of chemical that is either absorbed or available to be absorbed into the body where it can interact

with the target tissue (liver, thyroid, red blood cells, etc.). Knowledge of the exposure or concentration of a chemical is essential to determine the magnitude of the dose received. So, too, is knowledge of certain exposure factors such as the volume of contaminated air inhaled or food and water ingested per unit time. In fact, in many cases, it is quite simple to calculate dose when exposure concentration and exposure rate are known. To assist the risk assessor in making dose calculations, the U.S. EPA has published equations that are applicable to a variety of the most frequently encountered exposure scenarios. These equations, including the ones used below, are found in the document entitled *Risk Assessment Guidance for Superfund (RAGS), Human Health Evaluation Manual Part A (U.S. EPA, 1989)*. The following section provides example dose equations for some common exposure pathways:

Inhalation

$$\text{Dose}\,(\text{mg/m}^3) = \frac{\text{CA} \times \text{EF} \times \text{ED} \times \text{ET}}{\text{AT}}$$

where
CA = chemical concentration in air (mg/m^3),
EF = exposure frequency (days/year),
ED = exposure duration (years),
ET = exposure time (hours/day) × 1/24 day/hours, and
AT = averaging time (days).

Should it be desirable to express the safe air concentration in parts of toxicant per million parts of air, the dose (where the air concentration is in units of milligrams per cubic meter of air [mg/m^3]) may be converted to a ppm level by the following relationship:

$$\text{ppm} = \frac{\text{Dose}\,(\text{mg/m}^3) \times 24.5}{\text{MW}}$$

where MW is the molecular weight of the chemical (g/mol) and 24.5 is the amount (liters) of vapor per mole of contaminant at 25 °C and 760 mm Hg.

Ingestion of Groundwater

$$\text{Dose}\,(\text{mg/kg} \cdot \text{day}) = \frac{\text{CW} \times \text{IR} \times \text{EF} \times \text{ED}}{\text{BW} \times \text{AT}}$$

where
CW = chemical concentration in water (mg/l),
IR = ingestion rate (l/day),
EF = exposure frequency (days/year),
ED = exposure duration (years),
BW = body weight (kg), and
AT = averaging time (days).

Ingestion of Soil

$$\text{Dose}\,(\text{mg/kg} \cdot \text{day}) = \frac{\text{CS} \times \text{IR} \times \text{RBA} \times \text{EF} \times \text{ED}}{\text{BW} \times \text{AT}}$$

where
CS = chemical concentration in soil (mg/kg),
IR = ingestion rate (kg/day),
RBA = chemical-specific relative bioavailability,
EF = exposure frequency (days/year),
ED = exposure duration (years),
BW = body weight (kg), and
AT = averaging time (days).

Dermal Contact with Water

$$\text{Absorbed dose}\,(\text{mg/kg} \cdot \text{day}) = \frac{\text{CW} \times \text{SA} \times \text{PC} \times \text{ET} \times \text{EF} \times \text{ED} \times \text{CF}}{\text{BW} \times \text{AT}}$$

where
CW = chemical concentration in water (mg/l),
SA = skin surface area available for contact (cm^2),
PC = chemical-specific dermal permeability coefficient (cm/h),
ET = exposure time (hours/day),
EF = exposure frequency (days/year),
ED = exposure duration (years),
CF = conversion factor (1 l/1000 cm^3),
BW = body weight (kg), and
AT = averaging time (days).

Dermal Contact with Soil

$$\text{Absorbed dose}\,(\text{mg/kg} \cdot \text{day}) = \frac{\text{CW} \times \text{SA} \times \text{AF} \times \text{DA} \times \text{EF} \times \text{ED} \times \text{CF}}{\text{BW} \times \text{AT}}$$

where
CS = concentration in soil (mg/kg),
SA = skin surface area available for contact (cm^2/day),
AF = adherence factor for soil (mg/cm^2),
DA = dermal absorption,
EF = exposure frequency (days/year),
ED = exposure duration (years),
CF = conversion factor (10^{-6} kg/mg),
BW = body weight (kg), and
AT = averaging time (days).

To illustrate dose calculation, assume that a 16 kg child ingests 200 mg soil/day containing 400 mg/kg of chemical X, a volatile solvent. As shown in the following example, the child's dose of chemical X from the ingestion of soil is 1.25×10^{-3} mg/kg · day. This figure may not represent total dose, however, since dermal and/or inhalation exposure to the volatile chemical is likely. This illustrates the importance of considering all exposure pathways when assessing exposure

for the purpose of risk assessment. Failing to do so may result in the underestimation of risk.

The following formula can be used to determine the residential exposure from ingestion of a chemical in soil by a 5-year-old child receptor:

$$\begin{aligned}\text{Dose}\,(\text{mg/kg}\cdot\text{day}) &= \frac{\text{CS}\times\text{IR}\times\text{CF}\times\text{FI}\times\text{EF}\times\text{ED}}{\text{BW}\times\text{AT}}\\ &= \frac{400\text{mg/kg}\times 200\text{mg/day}\times 10^{-6}\,\text{kg/mg}\times 0.25\times 365\text{days/year}\times 6\text{years}}{16\text{kg}\times 2190\text{days}}\\ &= 1.25\times 10^{-3}\,\text{mg/kg}\cdot\text{day}\end{aligned}$$

where

CS = chemical concentration in soil (mg/kg) = 400 mg/kg (site-specific value),

IR = ingestion rate (mg soil/day) = 200 mg soil/day (default value for children 1–6 years old),

CF = conversion factor (10^{-6} kg/mg),

FI = fraction ingested from contaminated source (unitless) = 0.25 (site-specific value),

EF = exposure frequency (days/year) = 365 days/year (site-specific value),

ED = exposure duration (years) = 6 years (site-specific value),

BW = body weight (kg) = 16 kg (default value for children 1–6 years old), and

AT = averaging time (period over which exposure is averaged in days) = 6 years × 365 days/year = 2190 days for noncancer effects.

As shown in this example, where the intake of chemical X from soil ingestion was calculated for a 5-year-old child, default values for input variables can be used where site-specific data are lacking. The most complete collection of default values has been compiled and published by the U.S. EPA as the *Exposure Factors Handbook* (2011) and the *Child Exposure Factors Handbook* (2008). Numerous distributions (vs. discrete values) for input variables used in exposure calculations have also been reported that are of value to probabilistic risk assessment. In addition, the U.S. EPA has published several guidance documents that address many of the issues related to characterizing exposures for selected pathways:

- *Guidelines for Exposure Assessment* (USEPA, 1992)
- *Risk Assessment Guidance for Superfund* (USEPA, 1989)
- *A Framework for Assessing Health Risk of Environmental Exposures to Children* (USEPA, 2006)
- *Supplemental Guidance for Assessing Susceptibility from Early-Life Exposure to Carcinogens* (USEPA, 2005)
- *Exposure and Human Health Reassessment of 2,3,7,8-Tetrachlorodibenzo-p-Dioxin and Related Compounds, National Academy Sciences (External Review Draft)* (USEPA, 2004)
- *Guidance Document on the Development, Evaluation, and Application of Environmental Models* (USEPA, 2009)
- *Standard Scenarios for Estimating Exposure to Chemical Substances During Use of Consumer Products* (USEPA, 1986)
- *Available EPA Information on Assessing Exposure to Pesticides in Food: A User's Guide* (USEPA, 2000)
- *Standard Operating Procedures for Residential Pesticide Exposure Assessment* (USEPA, 2012)
- *Framework for Cumulative Risk Assessment* (USEPA, 2003)

Once dose has been estimated for all exposure pathways, it can be directly compared to toxicity values such as U.S. EPA RfDs and Agency for Toxic Substances and Disease Registry minimal risk levels (MRLs) to assess noncancer risks or, alternatively, multiplied by cancer slope factors to obtain an estimate of cancer risk. Another word of caution is in order, however. Toxicity values, including cancer slope factors, may be specific for particular exposure routes (MRLs vary by exposure route and exposure duration) since target organ dose and, for some chemicals, the target organ itself can be exposure route dependent.

While the administered dose and absorbed dose are common dose measures, they do not reflect the amount of the chemical or its metabolite(s) that ultimately produces the toxic response (except in cases where chemicals exert their action locally, as in the case of strong acids or bases that produce dermatotoxicity on contact). The toxic response is more closely linked to the dose in the target tissue of interest (see Figure 23.6 for a schematic showing the relationships between exposure and various dose measures). For example, solvent-induced neurobehavioral toxicity may be a function of peak brain concentration of the parent compound, whereas liver toxicity from the same chemical may be related to the hepatic tissue concentration of one or more metabolites over time (called the *area under the tissue concentration-time curve* (AUC)). The identification of such internal dose measures that are mechanistically linked to various toxicities holds promise for improving the risk assessment process. It is commonly assumed in interspecies extrapolation that the target tissue dose required to produce a biological effect of a given intensity is quantitatively similar across species. Therefore, dose–response curves generated with measures of target tissue dose should be more extrapolatable across

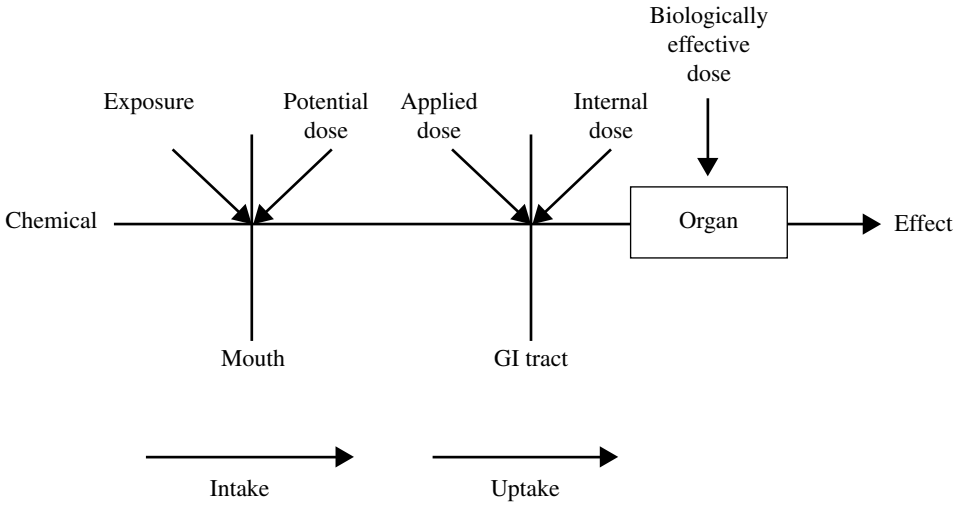


FIGURE 23.6 The relationship between exposure and various measures of dose. *Source:* Adapted from USEPA (1997).

species since they obviate the need for consideration of interspecies differences in the toxicokinetics of an administered dose. Unfortunately, dose–response curves of this nature are rare, primarily because of the technical difficulties inherent in internal dose measurement. This is likely to change, however, as advancements are made in analytical chemistry and PBPK models find their way into the mainstream. Such models are powerful tools with which to estimate internal dose measures from an endless variety of exposure scenarios to physiologically diverse receptors. As such, they are particularly valuable for the purpose of interspecies extrapolation (see Chapter 3).

Biological monitoring is another means of exposure assessment. When it is conducted to measure a chemical or its metabolites in the urine, blood, or tissue (including hair and fingernails) of an exposed individual, the chemical and its metabolites are referred to as *biomarkers of exposure*. Other potential biomarkers include DNA and protein adducts, mutations, chromosomal aberrations, genes that have undergone induction, and a host of other "early" cellular or subcellular events thought to link exposure and effect. The characterization and quantification of these latter biomarkers is known as *molecular dosimetry*. If found to be correlated with susceptibility, exposure, and effect, these biomarkers could considerably alter conventional approaches to risk assessment. Perhaps the best known example of such a correlation is urinary aflatoxin–DNA adducts and liver cancer. While molecular dosimetry holds promise for risk assessment, it is yet to be developed well enough for routine application.

Despite advancements in analytical chemistry, mathematical modeling, and biomonitoring, exposure assessment remains a challenge. It is important to realize that most exposure assessments result in estimates rather than definitive values. This stems in part from the fact that site- or situation-specific values are rarely available for all of the input variables necessary to calculate exposure. Despite the challenge, efforts should continue toward conducting exposure assessments that reflect realistic exposures. The identification of the dose metric that best correlates with various toxicities should also be a priority. Since this depends on a thorough knowledge of a chemical's mode of action, advancement in exposure assessment is inextricably linked to advances in toxicology.

23.5 RISK CHARACTERIZATION

The purpose of the risk characterization step is to integrate information provided by the hazard identification, dose–response assessment, and exposure assessment in order to develop risk estimates. Risk information may be conveyed in a qualitative manner, quantitative manner, or both. A qualitative assessment may describe the hazard posed by chemicals of concern, discuss opportunities for exposure, and reach some general conclusions that the risks are likely to be high or low, but would not provide numerical estimates of risk. A quantitative risk characterization, on the other hand, includes numerical risk values. Theoretically, there are many ways that numerical risks could be calculated depending on the specific questions being addressed in the risk assessment. For example, risks could be expressed as an individual's excess lifetime risk of developing a particular health effect as a result of chemical exposure. Risks could also be expressed on a population basis (e.g., estimated number of extra cases of a disease per year attributable to chemical exposure) or as the relative risk of an exposed population versus an unexposed population. Other ways of expressing risks, or health impacts, could be in terms of loss of life expectancy or lost days of work.

As discussed in Section 23.3, the most common means of expressing cancer risk associated with chemical exposure is in the form of individual excess lifetime cancer risk. When calculated for regulatory purposes, these values are intended to represent upper-bound estimates. That is, a cancer risk of one in one million means that an individual chosen at random from the exposed population is likely to have a probability no greater than one in one million of developing cancer as a result of that exposure. Note that this is an excess probability of developing cancer associated specifically with the chemical exposure addressed in the risk assessment, not the overall probability of developing cancer. The risk assessment provides an estimate of excess cancer risks, but does not determine whether the excess cancer risks are acceptable or unacceptable. That determination lies in the province of risk management, which must balance the risk estimate with other considerations (e.g., likelihood of actual exposure, uncertainties in the risk estimate, costs and feasibility of risk reduction strategies) to make decisions regarding steps, if any, to be taken to address chemical exposures. In some situations, an excess cancer risk of 1×10^{-3} (one in one thousand) from chemical exposure has been acceptable to regulatory agencies, while in others, any excess risk above 1×10^{-6} (one in one million) has been deemed too high. Travis et al. (1987) reviewed the risks associated with 132 federal regulatory decisions involving environmental carcinogens to determine the level of risk that led to regulatory action. Their analysis revealed that with large populations an action was always taken when the risk exceeded 10^{-4}. For small populations, historically, the *de manifestis level*, that is, the level at which action is always taken, was a risk of 10^{-3}. The *de minimis risk level* for these 132 regulatory actions, namely, the level of risk where no action or consideration is deemed necessary, was 10^{-5} to 10^{-4} for small populations and 10^{-6} to 10^{-7} for large populations. Others have suggested that the risk to smaller populations (e.g., a specific workforce) may be justifiably higher as long as the projected risk does not result in the expectation of an additional cancer. For example, if 100 persons were exposed to a 10^{-3} lifetime risk, the total population risk would be only 0.1, and an additional cancer case would be unlikely. The range of "acceptable risks" that has been applied by the U.S. EPA across its various regulatory programs seems to support the conclusions of the regulatory analysis performed by Travis and coworkers. The acceptable risk ranges of several U.S. EPA programs are:

- 10^{-4} to 10^{-6}: the cleanup policy under the U.S. EPA Superfund Cleanup Program of the National Oil and Hazardous Substances Pollution Contingency Plan
- 10^{-4} to 10^{-6}: U.S. EPA drinking water standards (MCLs) under the Safe Drinking Water Act
- 10^{-2} to 10^{-6}: National Emission Standards for Hazardous Air Pollutants (NESHAPs) under the Clean Air Act
- 10^{-4} to 10^{-6}: for corrective actions under the Resource Conservation and Recovery Act

For carcinogens with thresholds, risk is portrayed as a margin of exposure. (*Note:* The concept of margin of exposure is discussed in Section 23.3.) Just as the definition of an acceptable cancer risk in probability terms is outside the scope of the risk assessment, an acceptable margin of exposure is essentially a policy and risk management issue. The margin-of-exposure concept is also applicable to noncancer effects and is used to convey the difference between the estimated exposure to a chemical and the BMD, usually for the most sensitive effect. The other, more common, means of expressing hazard for noncancer effects is the HI. By convention, an HI greater than 1 signals concern for the possibility of adverse effects. The likelihood that health effects will actually occur with an HI greater than 1 depends in part on the chemicals in question and the margin of safety inherent in the toxicity values used to calculate the HI. These issues bear discussion in the risk characterization, so as to better inform the risk management decisions.

During the risk characterization step, risks from various chemicals, reaching individuals by various pathways and conceivably entering the body by various routes, must be combined in some way such that the total risk to individuals from chemical exposure can be assessed. Methods for combining risks are discussed in Section 23.7. It is not uncommon for risk estimates to be presented for a number of different populations. This may include groups of individuals exposed in different ways (e.g., workers at a contaminated site vs. visitors to the site vs. residents living nearby the site) or individuals that may differ in their sensitivity to hazards posed by the chemicals of concern (e.g., children, pregnant women). Development of these various risk estimates is important, not only in providing a complete characterization of potential risks posed by the chemicals but also in developing effective strategies for managing the risks.

A particularly important aspect of the risk characterization is a discussion of the uncertainties associated with the risk estimates. Each individual step in the risk assessment process is a potential source of uncertainty. Many of these are discussed throughout the chapter and include uncertainty associated with estimating exposure (e.g., measurement errors, uncertainty in selecting the best exposure models, uncertainty regarding exposure conditions that will exist in the future) as well as determining safe levels of exposure (e.g., uncertainty regarding the shape of the dose–response relationship in the low-dose region and extrapolating results from animals to humans, uncertainty that the most sensitive health effect has been identified, uncertainty regarding ways that multiple chemicals might interact). These need to be articulated in the risk characterization so that an appreciation of the level of confidence and conservatism in the risk estimate can be gained. Without a discussion of uncertainty, risk assessment results are often perceived as being more precise than they really are, which could lead to misuse. Minimally, uncertainties should be discussed qualitatively,

identifying the source or nature of each uncertainty and how, in a general way, it could affect the risk estimation (i.e., whether the approach taken, in view of the uncertainty, is likely to contribute to an over- or underestimation of risk). A semiquantitative perspective is helpful, in which the implications of model and assumption choices on the risk estimate are described in rough, order-of-magnitude terms. As discussed in Section 23.6, probabilistic techniques can be used to provide more precise quantitative expression of the uncertainties associated with risk estimates, as well as a description of variability in risks encountered in exposed populations. This approach is attractive in that it offers a much richer characterization of the risks and uncertainties than do the more traditional risk estimation techniques. It is technically demanding, however, requiring much greater time, resources, data, and technical expertise.

23.6 PROBABILISTIC VERSUS DETERMINISTIC RISK ASSESSMENTS

Risk assessments are rarely performed for a single individual. Instead, risk assessments are designed primarily to characterize risks to populations of individuals. Many—perhaps most—of the factors that affect risk can vary from one person to another. Differences in body weight, inhalation rate, frequency and duration of contact with contaminated media, and even sensitivity to toxicity are examples of factors that can lead to different risks among individuals, even if the concentration of chemical to which they are all exposed is the same. Theoretically, there is no single risk for a particular exposure circumstance, but rather as many different risk values as there are individuals in the exposed population. In confronting the issue of variability in risk assessment, the traditional approach used for regulatory purposes has been to simply characterize the risk to individuals within the population with the greatest exposure. A *deterministic approach* to risk calculation is used, where a single value is selected for each exposure variable and a single risk estimate is produced. The exposure assumptions are chosen to represent the plausible upper bound of exposure, and the risk estimate is said to be associated with reasonable maximal exposure (RME) or high-end exposure. The development of a single, high-end risk estimate for regulatory use is consistent with the goal of regulatory agencies to develop risk management strategies protective of the entire population. For perspective, a deterministic approach may also be used to develop an estimate of risk for the average individual in the population, that is, a central tendency estimate of the risk. The problem with this approach is that it provides little information on the extent to which risk varies within the population. For example, while exposure values may be selected to develop a high-end estimate of risk, it is difficult to know whether this value is merely conservative or extreme. Does it represent an exposure that might be exceeded by 1 or 2 out of every 10 individuals, or does it represent an exposure circumstance so extreme that it is unlikely ever to take place? This imprecision regarding the degree of conservatism in deterministic risk estimates undermines their value and creates controversy regarding their use in regulatory decision making.

A second problem confronting the risk assessor is management of uncertainty in the risk assessment process. As described elsewhere in this chapter, there are numerous sources of uncertainty in risk calculations, including uncertainty in the selection of models and assumptions and in measurements of risk-related parameters. As part of a deterministic calculation of risk, a choice must be made for each of these so that a risk estimate can be made. For regulatory purposes, conservative choices are usually made; models and assumptions that tend to provide higher estimates of risk are selected from among the range of plausible alternatives. The reason for conservative choices by regulatory agencies in the face of uncertainty is well understood, but the extent of conservatism imparted by the various choices is usually unclear. As with the issue of variability, this makes it difficult or impossible for the risk assessor to effectively convey the inherent conservatism associated with the risk estimate.

Probabilistic risk assessment is an alternative approach that can address the limitations of deterministic calculations in terms of variability and uncertainty. In probabilistic risk assessment, input variables are entered as probability density functions (PDFs) instead of single values. For example, instead of using a single body weight of 70 kg in the risk calculation, a distribution of body weights would be entered that reflects the variability in body weight of the exposed population. PDFs might also be entered for other variables such as inhalation rate, skin surface area, and frequency of contact with contaminated media—anything that would be expected to vary from one individual to another. These PDFs are then combined in such a way as to yield a risk distribution, representing the range and frequency of risks anticipated to exist in the exposed population. Although there are several ways to combine PDFs, one of the most commonly used techniques is Monte Carlo simulation. With Monte Carlo simulation, a computer program in essence creates a simulated population designed to resemble the exposed population in every key respect. For each risk calculation, it takes a value from each input PDF chosen in relation to its probability and calculates a numerical risk. This process is repeated, usually thousands of times, and the resulting range of risk values is tallied in the form of a distribution. This distribution represents the risk distribution for the population. From this distribution, the variability in risk among individuals can be visualized and the risk level at various percentiles of the population determined (see Figure 23.7).

Probabilistic risk assessment can also provide quantitative representation of the uncertainties in the risk calculation. For each input or model, some estimate of the uncertainty is

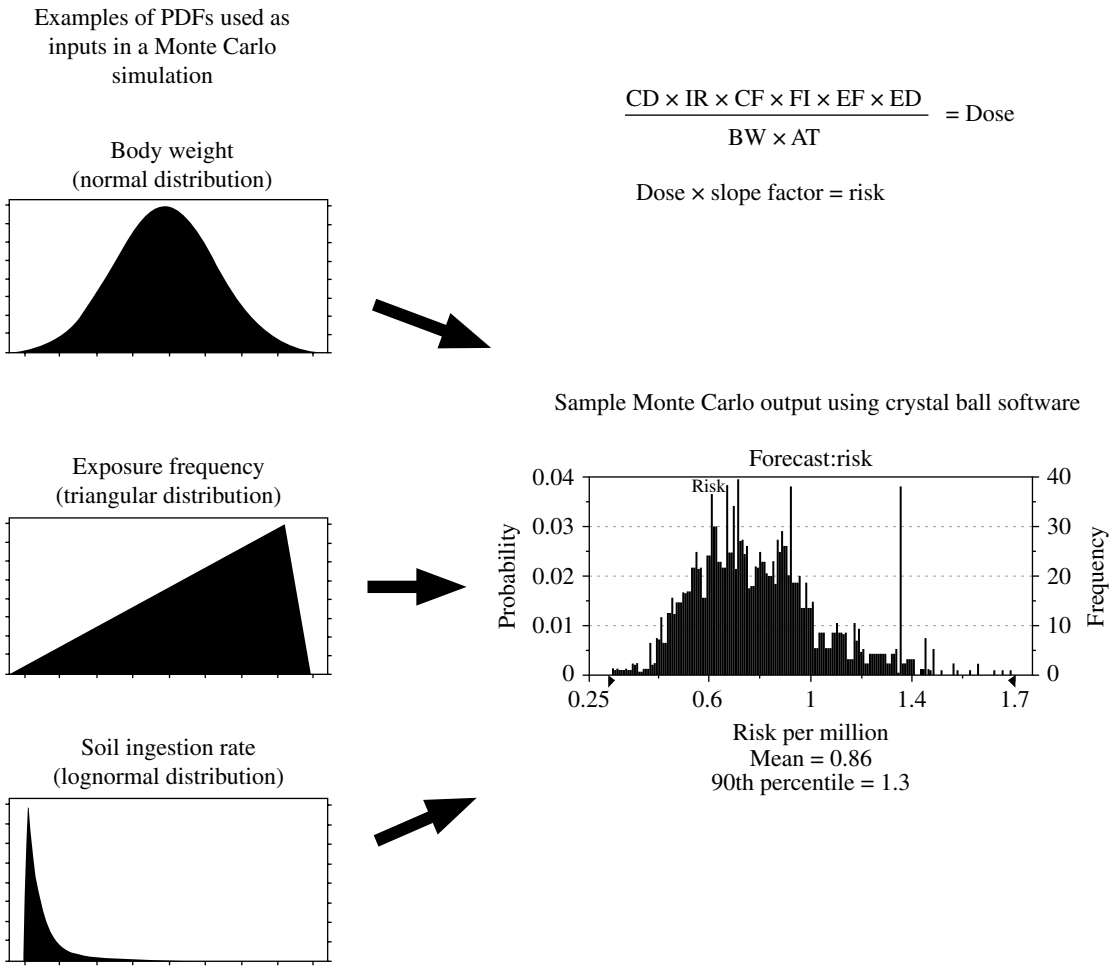


FIGURE 23.7 Simplified diagram of combining probability density functions (PDFs) to yield a risk distribution in probabilistic risk analysis. In this example, dose is calculated as described in Section 23.4, except each of the inputs for which variability exists is entered as a PDF. Examples of PDFs for body weight, exposure frequency, and soil ingestion rate are shown. Using Monte Carlo simulation, the computer takes values from these PDFs to calculate risk values for individuals in a simulated population. The distribution of these risk values is provided in the output of the simulation, as shown in the figure.

entered. For example, the concentration of chemical X for which a risk estimate is desired is assumed to be 100, but could be as low as 50 or as high as 200. In this case, the chemical concentration could be entered as a distribution of values, with 100 as the most likely estimate, but with a range extending from 50 to 200. As with variability, the uncertainty associated with various inputs can be combined to produce a PDF showing boundaries of uncertainty associated with a risk estimate. An additional benefit of this approach is that a sensitivity analysis can be used to rank the various sources of uncertainty in terms of their relative contribution to overall uncertainty. If the uncertainty is unacceptably large, a sensitivity analysis can be used to identify the best areas for further analysis or research to reduce uncertainty.

It is possible for a probabilistic risk assessment to address both variability and uncertainty simultaneously. This requires the development of PDFs for both uncertainty and variability. For example, a PDF might be used to portray variability in body weight in the exposed population, and a separate PDF would be used to deal with any uncertainty that the body weight distribution selected accurately reflects the actual body weight distribution of the population in question. (*Note:* This is not an unreasonable uncertainty, since risk assessors almost never have the time and resources to actually weigh everyone in an exposed population and therefore must rely on published body weight data for the general population to create their body weight PDF.) The variability and uncertainty PDFs are then analyzed separately to generate a risk distribution with confidence boundaries provided by the uncertainty distributions. This is called a *two-dimensional probabilistic risk assessment.*

The principal advantage of a probabilistic risk assessment is that it provides much greater information on variability and uncertainty associated with risk estimates. The manner

in which risk is distributed within the exposed population is transparent, and the magnitude of uncertainty associated with the risk estimate is conveyed in quantitative terms. There are, however, a number of disadvantages to probabilistic risk assessment, including the following: (i) it is technically demanding, requiring much greater expertise than deterministic risk calculations; (ii) it is information intensive, requiring data on exposure characteristics within populations that may not exist; (iii) because of the two previous points, it is much more time consuming and expensive than deterministic risk assessments; (iv) although they provide much more information, the outputs can be complex and difficult for nontechnical audiences to understand; and (v) it requires a different set of policy assumptions regarding acceptable risk. Unlike the situation with a single risk value, acceptable risk must be defined in terms of an acceptable risk distribution.

23.7 EVALUATING RISK FROM CHEMICAL MIXTURES

The simplest form of risk assessment deals with health risks posed by exposure to a single chemical from a single source. Unfortunately, in reality, things are seldom this simple. In most situations, exposure occurs not to a single chemical in high doses (as in toxicology studies in animals), but rather to multiple chemicals in lower doses. Often, more than one of these chemicals is capable of affecting the same target organ or tissue. In this situation, evaluating the risk for individual chemicals one by one may not accurately portray the risk associated with these chemicals in combination. In developing credible risk assessments, it is important not only to consider the cumulative impact of different chemicals affecting the same target organ but also to recognize the potential for these chemicals to interact. Effects of chemicals in combination may not be simply additive. Biological and chemical interactions among the chemicals can lead them to antagonize the effects of one another or produce effects greater than the sum of their individual effects. The ability to account for such interactions and develop meaningful estimates of risks of chemicals in combination is one of the most significant challenges in risk assessment.

There are three basic approaches to evaluating the toxic potential of chemicals in combination. If toxicity data are available for the specific chemical mixture of interest, a preferred approach is to treat the mixture as a single toxicological entity. That is, toxicological data from animals treated with the mixture can be used to identify an RfD, BMD, or slope factor for use in the risk assessment. For this approach to be valid, the chemical mixture to which individuals are exposed must be the same as the mixture used in the toxicity studies, not only in terms of the specific chemicals present but also their proportions. A second approach involves using toxicity data from a "sufficiently similar" mixture, if available, to develop a risk estimate. A similar mixture might, for example, have the same constituents but slightly different proportions, it might have several common components but lack one or two, or it might have one or more additional components. Similar mixtures would be expected to act by the same mechanism of action or produce the same type of toxicity. Beyond these general expectations, there are no firm criteria as to what constitutes "sufficiently similar," leaving this decision up to the judgment of the toxicologist or risk assessor.

If inadequate toxicity data are available for an identical or similar mixture, a third approach is to assess the toxicity of the mixture based on toxicity of its components. This last approach invariably requires assumptions regarding the presence and nature of chemical interactions. Interaction in this context means that one chemical alters the toxicity of one or more other chemicals in the mixture. The default assumption is usually no interaction among the chemicals; that is, in the absence of evidence to the contrary, the chemicals are assumed to act independently—each neither enhancing nor reducing the effect of the others. Chemicals that produce the same toxic effects are considered to act in an additive fashion in this situation, and the total risk is the sum of the risks posed by the individual component chemicals.

Although seemingly simple in concept, in practice, there are several ways to add the effects of chemicals. One way is to use *dose addition*, where the chemicals are considered functional clones of each other. This means that they produce the same toxic effects (or at least the same toxic effect of interest) through the same mode of action and have similar pharmacokinetic properties. These chemicals do not necessarily have identical dose–response curves, and in fact, there can be substantial differences in toxic potency. However, the relationships between the dose–response curves are such that differences in potency between chemicals can be represented by some constant proportion (e.g., one chemical might produce the same toxic response as another, but always at 1% of the dose). Experimentally, dose–response curves of such agents are parallel.

For groups of chemicals that fit this description, combined risks can be calculated using the *relative potency factor* (RPF) approach. One chemical in the mixture (usually the best characterized toxicologically) is designed as the *index chemical* and assigned an arbitrary potency factor of 1. Dose–response information for other chemicals is used to assign each a potency factor relative to the index chemical. For example, a chemical with a potency of 1/100th the index chemical would be assigned an RPF of 0.01, while a chemical 10 times as potent as the index chemical would have an RPF of 10. In the risk assessment, these RPF values are used to convert doses of the various chemicals in the group to toxicologically equivalent doses of the index chemical. These doses are then summed and used, along with a toxicity value for

the index chemical (e.g., RfD, slope factor, as appropriate) to derive a risk estimate for the group as a whole.

The RPF values for chemicals in the group may vary depending on the toxic effect of concern and perhaps the exposure circumstances. There are a few examples where all of the toxic effects of concern share a common mode of action and a single scaling factor is applicable for all effects and exposure conditions. This represents a special case of the RPF method termed the *toxic equivalency factor* (TEF) approach. An example of the use of the TEF approach is the risk assessment of polyhalogenated aromatic compounds. Most of the adverse health effects of concern for these compounds are thought to arise from a common mode of toxicity: Ah receptor activation. In this example, the index chemical is 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (2,3,7,8-TCDD), which is assigned a relative potency of 1. Based on studies of comparative potency in terms of Ah receptor-mediated toxicity, comparative potency factors (termed TEFs) have been determined for other polyhalogenated aromatics (e.g., PCDDs, PCDFs, PCBs). TEFs for PCB congeners relative to 2,3,7,8-TCDD are listed in Table 23.3. In assessing risks from exposure to a mixture of PCB congeners using the TEF approach, the TEF equivalents for all congeners present in the environmental sample are summed to derive a risk estimate for the PCB mixture.

Another means of adding chemical effects is the *HI* approach. This approach does not require the assumption of a common mode of toxicity, only that the chemicals share the same target organ or effect. In this approach, the dose of each chemical is compared with some representation of a threshold dose for toxicity. In practice, that may be an RfD or a BMD (see Section 23.3 for a discussion of reference and BMDs and their derivation). The dose for which a risk estimate is sought is divided by the threshold dose for that chemical in the target organ of interest and the result is termed the *HQ*. For example, if exposure to a chemical is predicted to result in a dose of 1 mg/kg · day, and the safe dose for the toxicity of concern is 10 mg/kg · day, the HQ is 1/10 or 0.1. HQ for each chemical affecting the target organ is then summed to obtain the HI. The interpretation of the magnitude of the HI is similar to that already discussed (see Section 23.3).

TABLE 23.3 Proposed Toxic Equivalency Factors (TEFs) for PCB Congeners Relative to Dioxin

IUPAC No.	Congener	TEF
	2,3,7,8-TCDD	1
77	3,3′,4,4′-TCB	0.0001
81	3,4,4′,5-TCB	0.0003
126	3,3′,4,4′,5-PeCB	0.1
169	3,3′,4,4′,5,5′-HxCB	0.03
105	2,3,3′,4,4′-PeCB	0.00003
114	2,3,4,4′,5-PeCB	0.00003
118	2,3′,4,4′,5-PeCB	0.00003
123	2′,3,4,4′,5-PeCB	0.00003
156	2,3,3′,4,4′,5-HxCB	0.00003
157	2,3,3′,4,4′,5′-HxCB	0.00003
167	2,3′,4,4′,5,5′-HxCB	0.00003
189	2,3,3′,4,4′,5,5′-HpCB	0.00003

Source: Adapted from Van den Berg et al. (2006).

Yet another way in which effects can be added is through *response addition*. This differs from dose addition methods in that the chemicals and their effects are assumed to be completely independent. For this approach, the percent of animals or humans expected to develop toxicity from each of the individual chemicals at their respective doses is estimated. These percentages are termed the "*responses.*" The probability that a toxic event will result from a combination of two chemicals can be expressed as follows:

$$R_{\text{both}} = 1 - (1 - R_{\text{chemical A}}) \times (1 - R_{\text{chemical B}})$$

When the probabilities are small, this reduces to simply

$$R_{\text{both}} = R_{\text{chemical A}} + R_{\text{chemical B}}$$

This approach is considered to be useful in summing a series of small component risks, but does not work well when one or more of the risks are large. In practice, response addition is used primarily in developing estimates of total cancer risks from more than one chemical or from chemical exposure by more than one route.

Each of the aforementioned approaches to combining risks assumes no interaction among chemicals. This is not always the case. It is possible that in some instances one chemical might antagonize or inhibit the toxicity of another. In this situation, the combination of chemicals would produce less-than-additive toxicity. This could conceivably occur through a variety of means depending on the mechanism(s) of toxicity of the chemicals and their toxicokinetics. Examples include effects to decrease toxicant absorption, increase its elimination or decrease its bioactivation, competition for receptor binding, or production of an opposing biochemical or physiological effect. Chemicals in combination can also produce greater-than-additive effects. When both chemicals are capable of producing the effect, this is termed *synergism*. The special case in which one of the two chemicals has no effect on its own, but nonetheless increases the toxicity of another, is termed *potentiation.*

There are a number of tests available to determine whether two chemicals interact in an additive, subadditive (i.e., antagonistic), or supra-additive (i.e., synergistic) fashion. One of the most straightforward is the construction of an isobologram (see Figure 23.8). A straight line is constructed on the graph representing dose pairs that would produce a specified effect level if there were no interaction between the chemicals. This line, an additivity line, varies in slope according to the relative potencies of each chemical. If the

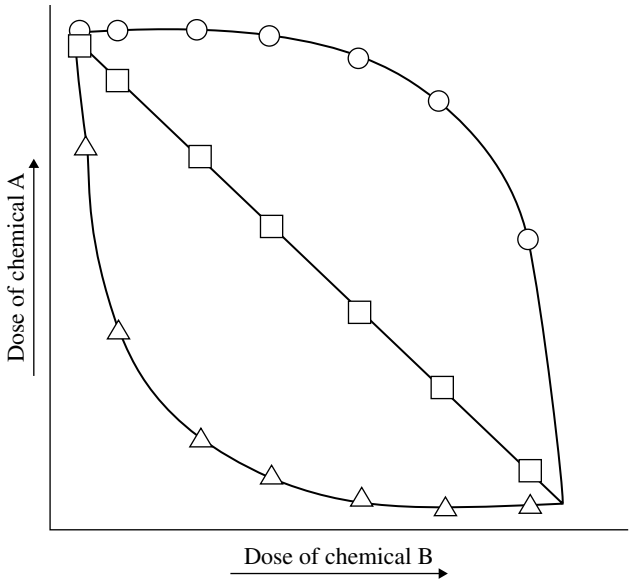


FIGURE 23.8 Isobologram of effects of two chemicals administered in varying dose combinations. The response obtained from chemical A is on the *y*-axis, and the response from chemical B is plotted on the *x*-axis. When there is no interaction between the chemicals, the responses from doses comprised of varying proportions of chemicals A and B will fall on the straight line (squares in the figure). If there is synergism between the chemicals, responses to combinations of the chemicals will lie below and to the left of this line (triangles), and antagonistic responses will lie above and to the right of the line (circles).

interaction between the two chemicals follows dose addition, their responses will lie along the additivity line. If the responses to chemical combinations are greater than would be predicted by dose addition, that is, if they lie below and to the left of the dose addition line, a synergistic effect can be inferred. This is because, in the presence of chemical A, a smaller dose of chemical B is needed to achieve the same effect. On the other hand, responses above the line and to the right indicate antagonism.

From a practical standpoint, interactions among chemicals are very difficult to deal with quantitatively in a risk assessment. These interactions are seldom well characterized and can be dose dependent such that synergism or antagonism that occurs at one dose combination of the chemicals may not occur at other dose combinations. Also, while tests exist to examine the nature of interactions between two chemicals, as described in the paragraph earlier, interactions among multiple chemicals are much more difficult to assess and characterize. Although the problem of addressing chemical interactions has been recognized for some time, research to solve this problem is still in a relatively early stage of development. Scientists are still struggling to identify circumstances where important interactions might take place, and rigorous techniques for adjusting risk estimates to account for interactions do not yet exist.

23.8 COMPARATIVE RISK ANALYSIS

For the purposes of this chapter, comparative risk analysis is a means of placing estimates of risk into a larger context in order to provide risk managers and stakeholders with a better perspective for decision making. Comparative risk analysis can also help nontechnical audiences understand the implications of a risk assessment, particularly when findings are reported in unfamiliar quantitative jargon. Furthermore, risk comparisons may be of value in setting priorities and allocating resources within regulatory agencies. In response to many risk problems posed by chemical exposure, the following questions might be asked, all of which would prompt a comparative risk analysis: (i) whether the receptors are exposed to the same chemical from other sources, (ii) whether exposure to the chemical also occurs from other environmental media, and (iii) whether other chemicals from the same sources pose additional risks to receptors. Several types of risk comparisons are listed below.

1. Comparisons of magnitude such as equating a "one-in-one-million" risk to the length of 1 inch in 16 miles, 30 s in a year, or 1 drop in 16 gallons
2. Comparisons of risk posed by the same chemical from different sources
3. Comparisons of risk posed by different chemicals from the same source
4. Comparisons of risk posed by different chemicals for the same target organ
5. Comparisons of familiar versus less familiar risks
6. Comparisons of voluntary versus involuntary risks
7. Comparisons of natural versus anthropogenic or technologic risks
8. Comparisons of risks of the same magnitude posed by different risk factors

Just as risk comparisons can be of value, they can also hinder risk communication. For example, inappropriate comparisons can be confusing and may serve to minimize risks that, in reality, deserve serious consideration. To maximize the benefits of risk comparison and avoid its pitfalls, it is recommended that substantially dissimilar risks (e.g., risk of cancer versus risk of losing money in the stock market) not be compared since the relative magnitudes of such risks are difficult to comprehend. Also, research on risk perception has suggested that directly comparing voluntary and involuntary risks or natural and technologic risks does not always improve a layperson's understanding of an environmental risk. However, the risk comparisons described in list items 2, 3, and 4 above are thought to be of communicative value.

There are no shortages of data available for risk comparisons, since we all incur risks by virtue of our continuous exposure to chemicals at work and at home. Indeed, the

TABLE 23.4 Cancer Risks for Indoor Air Exposures to VOCs

Chemical	Indoor Exposure Levels ($\mu g/m^3$)[a]	Potency $(\mu g/m^3)^{-1} \times 10^{-6}$	Lifetime Cancer Risk ($\times 10^{-6}$)
Benzene	15	3.3	50
Ethylbenzene	13	0.5	6
Methylene chloride	510	0.2	98
Tetrachloroethylene	7	1.2	8.6
Trichloroethylene	2.1	45	95
Chloroform	6.2	9.0	56
Vinyl chloride	0.04	2.5	0.1
Carbon tetrachloride	0.94	6.3	5.9

Source: Adapted from U.S. EPA (2011).

[a]Based on the maximum reported 90th percentile concentration measured in North American residences between 1990 and 2005.

TABLE 23.5 Cancer Risks for Household Exposures to Pesticides in the Lower Rio Grande Valley of Texas

Pesticide	Exposure (ng/m^3)[a]	Potency $(mg/kg \cdot day)^{-1}$	Lifetime Cancer Risk ($\times 10^{-6}$)
Heptachlor	1.6	525	840
Chlordane	1.8	42	75
Dieldrin	1.0	1890	1890
DDD	0.4	40	16
DDE	1.0	40	40
DDT	1.8	40	72
g-BHC (lindane)	11.6	150	1760
Simazine	0.4	14	5.6
Atrazine	3.2	27	86

Source: Adapted from Mukerjee et al. (1997).

[a]Based on the indoor median concentration detected in the spring.

TABLE 23.6 The Therapeutic and Virtually Safe Dosages (VSD) of a Few Medications

Drug	Cancer Slope Factor $(mg/kg \cdot day)^{-1}$	VSD $(mg/kg \cdot day)$	Dose Ratio (Daily Dose/VSD)	ILCR[a] per Daily Dose of Drug (10^{-6})[b]
Rifampin	2.1	4.8×10^{-7}	18,000,000	704.5
Isoniazid	4.9×10^{-1}	2.1×10^{-6}	2,400,000	93.9
Clofibrate	5.4×10^{-2}	1.9×10^{-5}	1,500,000	58.7
Disulfiram	3.6×10^{-1}	2.8×10^{-6}	1,300,000	50.9
Phenobarbital	2.9×10^{-1}	3.5×10^{-6}	825,000	32.3
Acetaminophen	1.3×10^{-2}	7.6×10^{-5}	747,000	29.2
Metronidazole	4.4×10^{-2}	2.3×10^{-5}	467,000	18.3
Sulfisoxazole	1.9×10^{-3}	5.3×10^{-4}	215,000	8.4
Dapsone	1.7×10^{-1}	5.8×10^{-6}	123,000	4.8
Methimazole	4.8×10^{-1}	2.1×10^{-6}	102,000	4.0
Oxazepam	1.0×10^{-1}	9.8×10^{-6}	87,700	3.4
Furosemide	6.0×10^{-2}	1.7×10^{-5}	68,723	2.7

Source: Adapted from Waddell (1996).

[a]ILCR—Incremental Lifetime Cancer Risk.

[b]Calculated by dividing Waddell's dose ration by the 25,550 days in a 70-year lifetime to get the incremental lifetime risk per daily dose of drug above the 10^{-6} risk representing the VSD.

potentially hazardous chemicals in the food we eat, the water we drink, and the air we breathe are numerous, and the list continues to grow as new studies are published. In addition, some medications carry a risk of cancer, and because the dosages of these chemicals are high relative to those chemicals found in the environment, over-the-counter medications and prescription drugs may carry significant theoretical risks even when used as intended. The following tables of risk comparisons have been provided to illustrate some different types of risk comparisons that can be made. Tables 23.4 and 23.5 illustrate the risks projected for volatile organic chemicals and pesticides measured in homes during the U.S. EPA study of residential environments (a type 3 risk comparison). Table 23.6 illustrates the theoretical risks associated with taking a daily dose of 12 different drugs (again, a type 3 risk comparison). Table 23.7 shows risk in a slightly different manner. In this table, risks for various activities, diseases, or lifestyle choices are compared by the number of days each is believed to decreases one's life expectancy (a comparison mixing categories 5, 6, and 7).

23.9 RISK COMMUNICATION

In order to be useful, risk assessment results must be effectively communicated to nontechnical audiences. This can include risk managers, legislators, the public, industry, and environmental groups. If risk managers do not understand the results, it can lead to bad regulatory and policy decisions. Public understanding of risk assessment results is also essential if they are to participate in and accept the results of risk-based decision making.

TABLE 23.7 Estimated Average Loss of Life Expectancy from Various Risks, Activities, and Diseases

	Days Lost
Alcohol addiction	4380
Poverty	3285
Smoking cigarettes and being male	2409
Heart disease	1606
Smoking cigarettes and being female	1424
Cancer	1247
Being an unmarried male	1168
Smoking cigars	1168
Being 35% overweight	964
Having less than an eighth-grade education	949
Being an unmarried female	694
Smoking a pipe	511
Stroke and being male	390
Being 25% overweight	303
Driving motor vehicle	207
Being a mine/quarry worker	167
Suicide	115
Pneumonia, influenza	105
Homicide	93
Misusing legal drugs	92
Diabetes	82
Accidents in the home	74
Falls	28
Suffocation	28
Oral contraceptives	25
Drowning	24
Generation of energy	24
Employment that entails radiation exposure	23
Fires and bums	20
Solid and liquid poisons	20
Dog bites	15
Natural radiation	10
Medical X-ray exposure	10
Firearm accidents	6.5
Riding a bicycle	6
Poisonous gases	4
Bee stings	0.2
Radiation from the nuclear industry	0.06

Source: Adapted from Cohen (1991, 2003).

Effectively communicating the results of risk assessments is an enormous challenge. Problems lie in virtually all aspects of the risk communication process, including (i) the individual, agency, or company that conducts and presents the risk assessment, (ii) the risk assessment itself, (iii) the means to convey risk information, and (iv) the audience. Examples of these problems are listed in Table 23.8. One of the biggest hurdles is the fact that risk analyses are often very complex, technical exercises. Making the process and outcome of the risk analysis transparent to laypersons is next to impossible unless there is some opportunity to provide background education to "bring them up to speed" on the subject. In most situations, this opportunity does not exist. The public is arguably one of the most important recipients of risk information yet one of the most difficult audiences for risk assessors to communicate with. One problem is that the most common channel for communicating risk information to the public is through the news media. This presents at least three difficulties in trying to communicate a clear and accurate message: (i) reporting of the information may be biased, incomplete, or inaccurate; (ii) news accounts may tend to sensationalize or focus on ancillary issues, such as disagreements between parties or human interest stories; and (iii) news media have generally shown little interest in providing the background information needed to educate the public on risk analysis and to help them interpret findings for themselves.

No doubt one reason why the media have not invested much effort in educating the public about risk assessment is that the public itself, for the most part, has shown little interest in the technical complexities and nuances of risk analysis. In most situations for which a risk assessment is needed, they just want a straight answer to the simple questions, "Is it safe?" Anything other than a clear "yes" answer to this question signals cause for concern. Herein lies a second major problem for risk communication. Unfortunately, all too often, the answers conveyed by the risk assessment can seem ambiguous. Scientists are trained to be circumspect in their conclusions and carefully point out any caveats in their analysis. This certainly applies to risk assessments, where responsible presentation of risk estimates is always accompanied by a discussion of the many areas of uncertainty and limitations in the analysis. When all of the caveats and uncertainties are presented along with the risk estimate, the uncertainty looms large and it is easy for the public to conclude "they don't really know what the risk is." When this happens, regardless of whether the risk estimates themselves are large or small, they have little credibility. Thus, the dilemma for the risk communicator is how to adequately convey the underlying uncertainties in the risk estimates without losing the essential message that the risks are large or small, as the case may be.

Deciding whether a risk is acceptable requires, in part, placing that risk in context. Thus, the risk from a particular chemical or set of exposure circumstances could be compared with other risks to the individual or population in order to place that risk in perspective. While this is straightforward in concept, it is difficult in practice, particularly when communicating risk to the general public. One reason is that the public, unaccustomed to seeing typical risk assessment outputs, may have little basis for comparison. Unless someone has experience with, or is shown, comparative risk data for a variety of hazards, it is difficult for them to know whether a 1×10^{-5} risk is significant. For noncancer health effects, the meaning of outputs in terms of HI or margin of exposure is even more obscure. How, for example, would you help citizens place an HI of 3 for a chemical exposure in the context of risk from events in their everyday lives?

TABLE 23.8 Examples of Risk Communication Problems

Source of Problem	Examples
Source of the message	The source of the risk information is not usually trained in communication skills
	The source of the risk information, usually a governmental or industrial entity or representative, is not trusted due to a history of exaggeration, secrecy, or worse
	Any disagreements among scientific experts make the information appear to be guesswork
	The risk assessment may not have received stakeholder input and, therefore, not address issues of greatest concern to individuals and communities
The message	Risk estimates may have large uncertainties due to limitations in data used in the risk assessment
	Risk assessments do not provide exact answers about the actual nature of the risk
Channel for conveying the message	Media interpretation may result in presentation of oversimplified, distorted, or erroneous information
	Media emphasis on drama, wrongdoing, or conflicts clouds presentation of risk information
	Journalists do not often have the scientific background needed to evaluate the disagreements or debates surrounding the risk estimates
Receiver of the message	Public perceptions of risk are often inaccurate
	There may be unrealistic demands for scientific certainty in risk estimates
	There may be a lack of interest in learning the technical complexities of the risk assessment and therefore a poor understanding of what risk estimates represent
	There may be difficulty in understanding risk if it relates to unfamiliar activities or technologies or is presented in an unfamiliar way
	Not everyone will be open minded; some individuals with strong opinions and beliefs will not be receptive to information that contradicts them
	Risks are usually perceived more in terms of outrange (or lack thereof) than actual harm or hazard

Source: Adapted from Covello and Sandman (2001).

A second reason that placing risks in context for the public is difficult is that the public often has distorted views of the risks posed by common and uncommon events in their lives. Comparing risks from chemical exposure to risks the public is more familiar is valuable only if their point of reference is accurate, and unfortunately, it seldom is. This has been demonstrated repeatedly in studies in which survey respondents' estimates of risks or comparative risk rankings for various hazards were compared with the actual, measured risks. Presenting the public with accurate risk comparisons can be helpful, but does not necessarily solve the problem. There are at least two reasons for this. One is that the meaning of the term "risk" itself is often different for the risk assessor and the public. The risk assessor tends to define risk as a probability of an adverse health effect and thinks of risk in purely probability terms. It is not surprising, then, that risk assessors once thought that a comparison of probabilities is all the public needs to place risks in perspective. The public, however, does not view risk simply in probability terms. The perception of the risk can be shaped powerfully by the nature of the risk (e.g., what health effect is at risk, such as cancer), whether the risk is voluntary or involuntary and whether the risk is accompanied by any perceived benefits.

Several strategies have evolved for improving risk communication. The first is to pay very careful attention to the language that is used in risk communication. Of course, jargon and acronyms unfamiliar to the public should be avoided. It is also important to understand that terms and expressions in common use in risk assessment have very different meanings to the public. For example, a "conservative approach" is understood in risk assessment to mean one protective of health, while the public might mistakenly interpret this as a risk assessment approach endorsed by one end of the political spectrum (e.g., as opposed to a "liberal approach"). In order to be more protective, an agency might "lower the standards" for a chemical, meaning to decrease permissible concentrations. To the public, however, lowering standards might be misinterpreted as allowing some sort of deterioration in their protectiveness. To avoid awkward and sometimes disastrous misunderstandings, it is important to carefully scrutinize the risk communication message and remove terms and phrases that will be unclear or have a different meaning for the public.

It is an unfortunate fact that there are few sources that the public explicitly trusts for risk information. Risk information provided by industry is often met with skepticism. In particular, risk messages that indicate no harm or basis for concern for chemical exposure are seen as self-serving. Credibility of governmental agencies charged with protecting public health and the environment is better, but not much. In dealing with the public, particularly when engaging them directly (e.g., through public meetings), it is extremely

important to be open and honest. An individual seen as not forthcoming with information, or who provides information solely as "technical gibberish," will be regarded as either completely out of touch or hiding something. From a risk communication standpoint, one is just as bad as the other. It is also important to listen to the public and gain an appreciation for their concerns and fears. Engaging in dialog early in the risk assessment process has several benefits, including the following:

1. It helps ensure that the risk assessment will be able to answer questions of greatest interest to the public.
2. Individuals in the public may be able to offer knowledge useful to the risk assessment, such as historical perspective and information regarding the manner in which individuals are (or have been) exposed to the chemicals in question.
3. It affords the opportunity to establish trust with the public. Of course, demeanor is important; a condescending manner is a sure way to cut the lines of risk communication.

23.10 SUMMARY

Conceptually, the basic components of any risk assessment are (i) hazard identification (what health effects may be produced by specific chemicals), (ii) dose–response assessment (what dose of chemical is required to produce these effects), (iii) exposure assessment (whether persons are actually exposed to chemicals and what doses they receive), and (iv) risk characterization (how likely is it that adverse effects will occur, and what are the potential limitations of the risk assessment as performed).

In order to fulfill their goal of ensuring protection of public health, regulatory agencies usually choose conservative exposure and modeling assumptions, namely, those that tend to overestimate rather than underestimate risk. Because the impact of each conservative assumption is frequently multiplicative and cumulative, the final risk estimate may overstate the true population risk substantially. Nonetheless, it is difficult to deviate from this approach given that considerable uncertainty exists for many components of the risk assessment.

While risk assessment has traditionally focused on human health, ecological risk assessments, which address potential impacts to plants and wildlife, are also commonly performed. Ecological risk assessments differ from human health risk assessments in that they are inherently more complex—there are many more species to consider, including interspecies relationships and more complicated exposure modeling—and they tend to focus more on population-, species-, and ecosystem-level effects.

Traditionally, cancer risks have been expressed in probability terms using linear, nonthreshold dose–response relationships. These relationships assume that any dose of a carcinogen poses some risk of developing cancer. The potential for noncancer health effects is evaluated using threshold models, where a dose below which no health effects will occur is assumed to exist. There has been increasing recognition that the dose–response relationship for some carcinogens may also involve a threshold, and methods to take this threshold into consideration in evaluating cancer risk from these chemicals have been used.

Deterministic risk assessments develop a single estimate of risk for a population, usually derived in such a way as to represent an upper-bound estimate. Probabilistic risk assessments can provide a description of the variability of risks within the population and quantitative estimates of uncertainty associated with those risks. While probabilistic risk assessments potentially offer more risk information, deterministic risk assessments are easier to perform and less expensive, and there exists a greater consensus as to how risk outputs should be conveyed and interpreted. At present, deterministic risk assessments are more routinely used because of their simplicity and ease of application.

The risk assessment should be performed in a *transparent* manner; that is, the steps performed should be easy to identify, understand, and evaluate. Also, the outcome of the risk assessment must be communicated in a way that can be understood by those without technical backgrounds, including the public. This is very challenging because risk assessors and the public may view risks and risk issues very differently.

A criticism of risk assessments that produce a numerical estimate of risk (quantitative risk assessments) is that they often convey the impression of greater precision than actually exists. It is vitally important that risk assessments include qualitative information as well, such as a discussion of the uncertainties associated with the risk estimate and the extent to which evidence of a true human hazard is weak or controversial.

It must be recognized that risk assessment is just one aspect of the larger process of risk management. In the development of strategies and procedures to address health concerns for chemical exposures, risk estimates undoubtedly play an important role. However, they are often not the sole consideration. Economic, social, and political factors, as well as technical feasibility, may also influence the management of chemical exposures in modern society.

SUGGESTED READING

Barnard RC. Scientific method and risk assessment. Regul Toxicol Pharmacol 1994;19:211–218.

Center for Risk Analysis. *A Historical Perspective on Risk Assessment in the Federal Government*. Boston: Harvard School of Public Health; 1994.

Cohen BL. Catalog of risks extended and updated. Health Phys 1991;61:317–335.

Cohen BL. Risks in perspective. J Am Phys Surg 2003;8(2):50–53.

Covello V, Sandman PM. Risk evolution and revolution. In: Anthony Wolbarst, editor. *Solutions to an Environment in Peril*. Baltimore: Johns Hopkins University Press; 2001. pp. 164–178.

Crump KS. The linearized multistage model and the future of quantitative risk assessment. Hum Exp Toxicol 1996;15:787–798.

Haddad S, Béliveau M, Tardif R, Krishnan K. A PBPK modeling-based approach to account for interactions in the health risk assessment of chemical mixtures. Toxicol Sci 2001;63(1): 125–131.

Mukerjee S, Ellenson WD, Lewis RG, Stevens RK, Somerville MC, Shadwick DS, Willis RD. An environmental scoping study in the Lower Rio Grande Valley of Texas—III. Residential microenvironmental monitoring for air, house dust, and soil. Environ Int 1997;23(5):657–673.

National Academy of Sciences (NAS). *Understanding Risk: Informing Decisions in a Democratic Society*. Washington: National Academy Press; 1996.

National Research Council (NRC). *Risk Assessment in the Federal Government: Managing the Process*. Committee on the Institutional Means for Assessment of Risks to Public Health, Commission on Life Sciences. Washington: National Academy Press; 1983.

National Research Council (NRC). *Science and Judgment in Risk Assessment*. Committee on Risk Assessment of Hazardous Air Pollutants, Board on Environmental Studies and Toxicology, Commission on Life Sciences. Washington: National Academy Press; 1994.

National Research Council (NRC). *Science and Decisions: Advancing Risk Assessment*. Committee on Improving Risk Analysis Approaches Used by the U.S. EPA, Board on Environmental Studies and Toxicology, Division on Earth and Life Sciences. Washington: National Academy Press; 2009.

Office of Technology Assessment. *Assessment of Technologies for Determining Cancer Risks from the Environment*. Washington: US Government Printing Office; 1981.

Sandman PM. *Responding to Community Outrage: Strategies for Effective Risk Communication*. Fairfax: AIHA Press; 1993.

Travis CC, Richter SA, Crouch EA, Wilson R, Klema ED. Cancer risk management. A review of 132 federal regulatory decisions. Environ Sci Technol 1987;21(5):415–420.

Travis CC, Hester ST. Background exposure to chemicals: what is the risk? Risk Anal 1990;10:463–466.

Tsuji JS, Yost LJ, Barraj LM, Scrafford CG, Mink PJ. Use of background inorganic arsenic exposures to provide perspective on risk assessment results. Regul Toxicol Pharmacol 2007; 48(1):59–68.

United States Environmental Protection Agency (USEPA). *Standard Scenarios for Estimating Exposure to Chemical Substances During Use of Consumer Products, Volume I*. Washington: Office of Toxic Substances, Exposure Evaluation Division; 1986.

United States Environmental Protection Agency (USEPA). *Risk Assessment Guidance for Superfund, Vol. I; Human Health Evaluation Manual, Part A, Interim Final*. Washington: Office of Emergency and Remedial Response; 1989.

United States Environmental Protection Agency (USEPA). *Guidelines for Exposure Assessment*. Washington: Risk Assessment Forum; 1992.

United States Environmental Protection Agency (USEPA). *Guidance Manual for the Integrated Exposure Uptake Biokinetic Model for Lead in Children*. Washington: OSWER. #9285.7-15-1; 1994.

United States Environmental Protection Agency (USEPA). *Exposure Factors Handbook*. Washington: National Center for Environmental Assessment; 1997.

United States Environmental Protection Agency (USEPA). *Guidelines for Ecological Risk Assessment*. Washington: Risk Assessment Forum. EPA/630/R-95/002F; 1998.

United States Environmental Protection Agency (USEPA). *Available Information on Assessing Exposure from Pesticides in Food: A User's Guide*. Washington: Office of Pesticide Programs; 2000.

United States Environmental Protection Agency (USEPA). *Risk Assessment Guidance for Superfund: Volume III—Part A, Process for Conducting Probabilistic Risk Assessment*. Washington: Office of Emergency and Remedial Response. EPA/540/R-02/002; 2001.

United States Environmental Protection Agency (USEPA). *A Review of the Reference Dose and Reference Concentration Process*. Risk Assessment Forum, Washington, DC. EPA/630/P-02/002F; 2002.

United States Environmental Protection Agency (USEPA). *Framework for Cumulative Risk Assessment*. Washington: Risk Assessment Forum; 2003.

United States Environmental Protection Agency (USEPA). *Risk Assessment Guidance for Superfund, Volume I: Human Health Evaluation Manual (Part E, Supplemental Guidance for Dermal Risk Assessment)*. Washington: Office of Superfund Remediation and Technology Innovation. EPA/540/R/99/005; 2004.

United States Environmental Protection Agency (USEPA). *Guidelines for Carcinogen Risk Assessment*. Risk Assessment Forum, Washington, DC. EPA/630/P-03/001F; 2005.

United States Environmental Protection Agency (USEPA). *Approaches for the Application of Physiologically Based Pharmacokinetic (PBPK) Models and Supporting Data in Risk Assessment*. Washington: National Center for Environmental Assessment, Office of Research and Development. EPA/600/R-05/043F; 2006.

United States Environmental Protection Agency (USEPA). *Framework for Metals Risk Assessment*. Washington: Office of the Science Advisor. Risk Assessment Forum. EPA/120/R-07/001; 2007.

United States Environmental Protection Agency (USEPA). *Framework for Application of the Toxicity Equivalence Methodology for Polychlorinated Dioxins, Furans, and Biphenyls in Ecological Risk Assessment*. Washington: Office of the Science Advisor. Risk Assessment Forum. EPA/100/R-08/004; 2008.

United States Environmental Protection Agency (USEPA). *Risk Assessment Guidance for Superfund, Volume I: Human Health Evaluation Manual (Part F, Supplemental Guidance for Inhalation Risk Assessment)*. Washington: Office of Superfund Remediation and Technology Innovation. EPA/540/R/070/002; 2009.

United States Environmental Protection Agency (USEPA). *Background Indoor Air Concentrations of Volatile Organic Compounds in North American Residences (1990–2005): A Compilation of Statistics for Assessing Vapor Intrusion*. Washington: Office of Solid Waste and Emergency Response. EPA/530/R/10/001; 2011.

United States Environmental Protection Agency (USEPA). *Benchmark Dose Technical Guidance*. Risk Assessment Forum, Washington, DC. EPA/100/R-12/001; 2012.

Van den Berg M, Birnbaum LS, Denison M, De Vito M, Farland W, Feeley M, Fiedler H, Hakansson H, Hanberg A, Haws L, Rose M, Safe S, Schrenk D, Tohyama C, Tritscher A, Tuomisto J, Tysklind M, Walker N, Peterson RE. The 2005 World Health Organization reevaluation of human and mammalian toxic equivalency factors for dioxins and dioxin-like compounds. Toxicol Sci 2006;93(2):223–241.

Waddell WJ. Reality versus extrapolation: an academic perspective of cancer risk regulation. Drug Metab Rev 1996;28(1–2): 181–195.

White RH, Cote I, Zeise L, Fox M, Dominici F, Burke TA, White PD, Hattis DB, Samet JM. State-of-the-science workshop report: issues and approaches in low-dose–response extrapolation for environmental health risk assessment. Environ Health Perspect 2009;117(2):283–287.

GLOSSARY

Absorption. The movement of a chemical from the site of initial contact with the biologic system across a biologic barrier and into either the bloodstream or the lymphatic system.

Accumulative effect of a chemical. The effect of a chemical on a biologic system when the chemical has been administered at a rate that exceeds its elimination from the system. Sufficient accumulation of the chemical in the system can lead to toxicity.

Acetylation. The introduction of an acetyl group, $CH_3CO–$, onto the molecule of an organic compound having either –OH or $–NH_2$ groups.

Acetylator. An individual with a phenotype of rapid metabolic acetylation; common in American Indians and those of Asian descent.

Acetylcholine. An acetic acid ester of choline normally present in many parts of the body and having important physiologic functions, such as playing a role in the transmission of an impulse from one nerve fiber to another across a synaptic junction.

Acetylcholinesterase. An enzyme present in the nervous tissue and muscle that catalyzes the hydrolysis of acetylcholine to choline and acetic acid.

Acidosis. A pathologic condition resulting from the accumulation of acid in, or loss of base from, the body.

Action potential. A momentary change in electrical potential on the surface or a nerve or muscle cell that takes place when it is stimulated, especially by the transmission of a nerve impulse.

Acutely toxic. Adverse effects caused by a toxic agent and occurring within a short period of time following exposure.

Adduct. A chemical addition product (i.e., a chemical bound to an important cellular macromolecule like DNA or protein).

Adenocarcinoma. A malignant tumor originating in the glandular tissue.

Adenoma A benign epithelial tumor having a glandular origin and structure.

Administrative control. A method of controlling employee exposures to contaminants by job rotation or work assignment within a single work shift.

Aflatoxins. Toxic metabolites produced by some strains of the fungus *Aspergillus flavus.* They are widely distributed in foodstuffs, especially peanut meals.

Albuminuria. Presence of serum albumin in the urine; proteinuria.

Alcohol. An organic compound in which a hydrogen atom attached to a carbon atom in a hydrocarbon is replaced by a hydroxyl group (OH). Depending on the environment of the –C–OH grouping, they may be classified as primary, secondary, or tertiary alcohols.

Aldehyde. A broad class of organic compounds having the generic formula RCHO.

Alicyclic. Organic compounds characterized by arrangement of the carbon atoms in closed-ring structures.

Aliphatic. Organic compounds characterized by a straight- or branched-chain arrangement of the constituent carbon atoms.

Alkane. See **Paraffin**.

Alkyl. A chemical group obtained by removing a hydrogen atom from an alkane or other aliphatic hydrocarbons.

Alkylation. The introduction of one or more alkyl radicals (e.g., methyl, $CH_3–$; ethyl, $C_2H_5–$; propyl,

Principles of Toxicology: Environmental and Industrial Applications, Third Edition. Edited by Stephen M. Roberts, Robert C. James and Phillip L. Williams.

$CH_3CH_2CH_2$–; etc.) by addition or substitution into an organic compound.

Allele. Either of the pair of alternative characters or genes found at a designated locus on a chromosome. Chromosome pairing results in expression of a single allele at each locus.

Allergy. General or local hypersensitive reactions of body tissues of certain persons to certain substances (allergens) that, in similar amounts and circumstances, are innocuous to other persons. Allergens can affect the skin (producing urticaria), the respiratory tract (asthma), or the gastrointestinal tract (vomiting and nausea) or may result from injections into the bloodstream (anaphylactic reaction). See also **Anaphylactic-type reaction.**

Alveolar macrophages. Actively mobile, phagocytic cells that process particles ingested into the lung. They originate outside the lungs from precursor cells (promonocytes) in the bone marrow and from peripheral blood monocytes. They enter the alveolar interstices from the bloodstream and are able to migrate to terminal bronchioles and lymphatic vessels.

Alveolus (pl. alveoli). In the lungs, small outpouchings along the walls of the alveolar sacs, alveolar ducts, and terminal bronchioles, through the walls of which gas exchange takes place between alveolar air and pulmonary capillary blood.

Amelia. The congenital absence of a limb or limbs. See also **Phocomelia**.

Amidase. An enzyme that catalyzes the breakdown of an amide compound to a carboxylic acid and ammonia.

Ames assay. A screening test capable of revealing mutagenic activity through reverse mutation in *Salmonella typhimurium.* Mammalian metabolism can be simulated by addition of S9 liver enzyme to the bacterial growth medium.

Amide. A nitrogenous compound with the general formula $RNH_2C=O$, related to or derived from ammonia. Reaction of an alkali metal with ammonia yields inorganic amides (e.g., sodium amide, $NaNH_2$–). Organic amides are closely related to organic acids and are often characterized by the substitution of one or more acyl groups (RCO) for an H atom of the ammonia molecule (NH_3).

Amine. An organic compound formed from ammonia (NH_3) by replacement of one or more of the H atoms by hydrocarbon radicals.

Amyotrophic lateral sclerosis (ALS). A disease marked by progressive degeneration of the neurons that give rise to the corticospinal tract and of the motor cells of the brainstem and spinal cord, resulting in a deficit of upper and lower motor neurons; the disease is usually fatal within 2–3 years.

Anaphylactic-type reaction. One of the four types of allergic reaction. A violent allergic reaction to a second dose of a foreign protein or other antigens to which the body has previously been hypersensitized. Symptoms include severe vasodilation, urticaria or edema, choking, shock, and loss of consciousness. Can be fatal.

Angiosarcoma. Malignant tumor of vascular system arising from endothelial cells.

Anoxia. A complete reduction in the oxygen concentration supplied to cells or tissues.

Anthropogenic. Produced or caused by the actions of humans.

Antibody. An immunoglobulin molecule that has a specific amino acid sequence that causes it to interact only with the antigen that induced its synthesis or with antigens closely related to it.

Antigen. A substance that, when introduced into the body, is capable of inducing the formation of antibodies and, subsequently, of reacting in a recognizable fashion with the specific induced antibodies.

Antipyretic. An agent that relieves or reduces fever.

Aplasia. Lack of development of an organ or tissue or of the cellular products of an organ or tissue.

Aplastic anemia. A form of anemia generally unresponsive to specific antianemia therapy, in which the bone marrow may not necessarily be acellular or hypoplastic but fails to produce adequate numbers of peripheral blood elements; term is all-inclusive and probably encompasses several chemical syndromes.

Apnea. Cessation of breathing; asphyxia.

Aromatic. A major group of unsaturated cyclic hydrocarbons containing one or more rings. These are typified by benzene, which has a six-carbon ring containing three double bonds. These are also known as *arene compounds.*

Arrhythmia. Any variation from the normal rhythm of heartbeat, including sinus arrhythmia, premature beat, heart block, atrial fibrillation, atrial flutter, pulsus alternans, and paroxysmal tachycardia.

Arteriosclerosis. A disease of the arteries characterized by thickening, loss of elasticity, and calcification of arterial walls, resulting in a decreased blood supply particularly to the cerebrum and lower extremities; it often develops with aging and in hypertension and diabetes.

Arthralgia. Neuralgic pain in a joint or joints.

Arthroosteolysis. Dissolution of the bone; the term is applied especially to the removal or loss of calcium from the bone; the condition is attributable to the action of phagocytic kinds of cells.

Asbestosis. A bilateral, diffuse, interstitial pulmonary fibrosis caused by fibrous dust of the mineral asbestos; also referred to as asbestos pneumoconiosis.

Asphyxiant. A substance capable of producing a lack of oxygen in respired air, resulting in pending or actual cessation of apparent life.

Asthmatic response. Condition marked by recurrent attacks of paroxysmal dyspnea, with wheezing caused by spasmodic contractions of the bronchi; the response is a reaction in sensitized persons.

Ataxia. Failure of muscular coordination; irregularity of muscular action.

Atherosclerosis. A form of arteriosclerosis characterized by the deposition of atheromatous plaques containing cholesterol and lipids on the innermost layer of the walls of large- and medium-sized arteries.

Atresia. A congenital absence of closure of a normal body orifice or tubular organ; the absence or closure of a normal body orifice or tubular passage such as the anus, intestine, or external ear canal.

Atrophy. A decrease in the size and activity of cells, resulting from such factors as hypoxia, decreased work, and decreased hormonal stimulation.

Atropine. An alkaloid forming white crystals, $C_{17}H_{23}NO_{31}$, soluble in alcohol and glycerin. Used as an anticholinergic for relaxation of smooth muscles in various organs to increase the heart rate by blocking the vagus nerve.

Autonomic nervous system. The part of the nervous system that regulates the activity of cardiac muscle, smooth muscle, and glands.

Auxotroph. Any organism (e.g., a bacterium) that, as a result of mutation, can no longer synthesize a substance that is necessary for its own nutrition (usually an amino acid) and thus requires an external supply of that substance.

Autosome. Any chromosome that is not a sex chromosome.

B cell. An inununocyte produced in the bone marrow. B cells are responsible for the production of immunoglobulins but do not play a role in cell-mediated immunity. They are short-lived.

Bactericidal. Destructive to bacteria.

Basophil. A granular leukocyte with an irregularly shaped, relatively pale-staining nucleus that is partially constricted into two lobes; cytoplasm contains coarse, bluish black granules of variable size.

Benign tumor. A new tissue growth (tumor) composed of cells that, although proliferating in an abnormal manner, are not invasive—that is, do not spread to surrounding, normal tissue; benign tumors are contained within fibrous enclosures.

Bilirubin. A bile pigment; it is a breakdown product of heme formed from the degradation of erythrocyte hemoglobin in reticuloendothelial cells, but is also formed by the breakdown of other heme pigments. Normally, bilirubin circulates in plasma as a complex with albumin and is taken up by the liver cells and conjugated to form bilirubin deglucuronide, which is the water-soluble pigment excreted in bile.

Biologic half-life. The time required to eliminate one-half of the quantity of a particular chemical that is in the system at the time the measurement is begun.

Biomarker. A specific physical or biochemical trait used to measure or indicate the effects or progress of a disease or condition.

Biotransformation. The series of chemical alterations of a foreign compound that occur within the body, as by enzymatic action. Some biotransformations result in less toxic products, while others result in products more toxic than the parent compound.

Bradycardia. A slowness of the heartbeat, as evidenced by a slowing of the pulse rate to less than 60 beats per minute.

Bronchitis. Inflammation of one or more bronchi, the larger air passages of the lungs.

Byssinosis. Respiratory symptoms resulting from exposure to the dust of cotton, flax, and soft hemp. Symptoms range from acute dyspnea with cough and reversible breathlessness and chest tightness on one or more days of a workweek to permanent respiratory disability owing to irreversible obstruction of air passages.

Cancer. A process in which cells undergo some change that renders them abnormal. They begin a phase of uncontrolled growth and spread. See also **Malignant tumor**.

Carbamate. A compound based on carbamic acid, NH_2COOH, which is used only in the form of its numerous derivatives and salts; as pesticides, carbamates are reversible inhibitors of cholinesterase. Inhibition of the enzyme is reversed largely by hydrolysis of the carbamylated enzyme and to a lesser extent by synthesis of a new enzyme.

Carcinogen. Any cancer-producing substance.

Carcinoma. A malignant tumor that arises from embryonic ectodermal or endodermal tissue.

Cathartic. An agent that stimulates bowel movement; a strong laxative.

Cardiomyopathy. General diagnostic term designating primary myocardial disease, often of obscure or unknown etiology.

Catecholamine. Any of a group of amines derived from catechol that have important physiological effects as neurotransmitters and hormones and include epinephrine, norepinephrine, and dopamine.

Cell-mediated immunity. Specific acquired immunity in which the role of small lymphocytes of thymic origin is predominant; the kind of immunity that is responsible for resistance to infectious diseases caused by certain bacteria and viruses, certain aspects of resistance to cancer, delayed hypersensitivity reactions, certain autoimmune disease, and allograft rejections and that plays a part in certain allergies. See also **T cell.**

Cellular immunity. Immunity mediated by T lymphocytes. It can be transferred to a naïve individual with cells but not by serum or plasma.

Cephalosporidine. A broad-spectrum antibiotic of the cephalosporin group, which are penicillinase-resistant antibiotics.

Chelate. A chemical compound in which a metallic ion is sequestered and firmly bound with the chelating molecule; used in chemotherapeutic treatments for metal poisoning.

Chemotaxis. The characteristic movement or orientation of an organism or cell along a chemical concentration gradient either toward or away from the chemical stimulus.

Chloracne. An acne-like skin eruption caused by exposure to halogenated compounds, especially the polyhalogenated naphthalenes, biphenyls, dibenzofurans, and dioxins.

Cholestasis. Stoppage or suppression of the flow of bile.

Cholestatic. Pertaining to or characterized by cholestasis.

Cholinesterase. An enzyme found chiefly at nerve terminals that inactivates the neurotransmitter acetylcholine by hydrolyzing it to form acetic acid and choline.

Chromhidrosis. The secretion of colored sweat.

Chromophore. A chemical group capable of selective light absorption resulting in the coloration of certain organic compounds.

Chromosome aberrations. Structural mutations (breaks and rearrangements of chromosomes) or changes in number of chromosomes (additions and deletions).

Chronic toxicity. Adverse effects occurring after a long period of exposure to a toxic agent (with animal testing, this is considered to be the majority of the animal's life). These effects may be permanent or irreversible.

Clastogenic. Giving rise to or inducing a breakage or disruption of chromosomes.

Cocarcinogen. Any chemical capable of increasing the observed incidence of cancer if applied with a carcinogen, but not itself carcinogenic.

Collagen. The fibrous protein constituent of the bone, cartilage, tendon, and other connective tissues. It is converted into gelatin by boiling.

Comedo. A plug in an excretory duct of the skin, containing microorganisms and desquamated keratin; a blackhead.

Competitive inhibition. Inhibition of enzyme activity in which the inhibitor (substrate analog) competes with the substrate for binding sites on the enzymes; such inhibition is reversible since it can be overcome by increasing the substrate concentration.

Complement. A complex system of proteins found in normal blood serum that combines with antibodies to destroy pathogenic bacteria and other foreign cells.

Conformational change. A change in the particular shape of a molecule.

Conjugation. The addition by drug-metabolizing enzymes of hydrophilic moieties to xenobiotics to hasten their excretion from the body.

Conjunctivitis. Inflammation of the conjunctiva, the delicate mucous membrane that lines the eyelids and covers the exposed surface of the eye.

Contact dermatitis. See **Dermatitis**.

Contraindication. Any condition, especially one of disease, that renders some particular line of treatment improper or undesirable.

Corpus luteum. A yellow glandular mass in the ovary formed by an ovarian follicle that has matured and discharged its ovum.

Cryptorchidism. A development defect marked by the failure of the testes to descend into the scrotum.

Cutaneous sensitization. Immune reaction characterized by local skin rashes, urticaria (hives), erythema, edema, and itching. Cutaneous sensitization is thought to be initiated by the release of histamine.

Cyanosis. A bluish discoloration, especially of the skin and mucous membranes, owing to excessive concentration of reduced hemoglobin in the blood.

Cystitis. Inflammation of the urinary bladder.

Cytochrome oxidase. An oxidizing enzyme containing iron and a porphyrin, found in mitochondria and important in cellular respiration as an agent of electron transfer from certain cytochrome molecules to oxygen molecules.

Cytochrome enzymes. See **Mixed-function oxidase system (MFO)**.

Cytokinesis. The division of the cytoplasm of a cell following the division of the nucleus.

Cytoplasm. The protoplasm of a cell exclusive of the nucleus, consisting of a continuous aqueous solution (cytosol) and the organelles and inclusions suspended in it (phaneroplasm); the site of most of the chemical activities of the cell.

Cytosol. The liquid medium of the cytoplasm (i.e., cytoplasm minus organelles and nonmembranous insoluble components).

Dalton. A unit of mass, one-twelfth the mass of the carbon-12 atom. Carbon-12 has a mass of 12.011, and thus, the dalton is equivalent to 1.0009 mass units, or 1.66×10^{-24} g. Also called the atomic mass unit (amu).

Denaturation. The destruction of the usual nature of a substance, usually the change in the physical properties of proteins caused by heat or certain chemicals.

Depolarize. Loss of the ionic gradient across a nerve cell membrane, resulting in an action potential and propagation of a nerve impulse.

Dermatitis. Inflammation of the skin. Contact dermatitis is a delayed allergic skin reaction resulting from contact with an allergen. Irritant dermatitis describes irritation of the skin accompanying exposure to a toxic substance.

Detoxification. The metabolic process by which the toxic qualities of a poison or toxin are reduced by the body.

Diethylstilbestrol (DES). A synthetic estrogenic compound, $C_{18}H_{20}O_2$, prepared as a white odorless crystalline powder.

Dimethyl sulfoxide (DMSO). An alkyl sulfoxide, C_2H_6OS, practically colorless in its purified form. As a highly polar organic liquid, it is a powerful solvent, dissolving most aromatic and unsaturated hydrocarbons, organic compounds, and many other substances.

Diplopia. A condition in which a single object is perceived as two objects; double vision.

Direct carcinogen. See **Primary carcinogen**.

Dissociation constant. The equilibrium constant for the reaction by which a weak acid compound is dissociated into hydrogen ions and a conjugate base, in solution. See also **p*K***.

Distal alveolar region. The part of the lung composed of the alveoli, or tiny air sacs, through which gas exchange between alveolar air and blood takes place.

DMSO. See **Dimethyl sulfoxide**.

Dose. The amount of a drug needed at a given time to produce a particular biologic effect. In toxicity studies, it is the quantity of a chemical administered to experimental animals at specific time intervals. The quantity can be further defined in terms of quantity per unit weight or per body surface area of the test animal. Sometimes, the interval of time over which the dose is administered is part of the dose terminology. Examples are grams (or milligrams) per kilogram of body weight (or per square meter of body surface area).

Dose–response relationship. One of the most basic principles of both pharmacology and toxicology. It states that the intensity of responses elicited by a chemical is a function of the administered dose (i.e., a larger dose produces a greater effect than a smaller dose, up to the limit of the capacity of the biologic system to respond).

Drug-induced toxicity. Toxicities that are "side effects" to the intended beneficial effect of a drug. They represent pharmacologic effects that are undesirable but that are known to accompany therapeutic doses of the drug.

Dyscrasia. A morbid general state resulting from the presence of abnormal material in the blood.

Dysplasia. Abnormal development or growth of tissues, organs, or cells.

Dyspnea. Difficult or labored breathing.

Dysrhythmia. Disturbances of rhythm, such as speech, brain waves, and heartbeat.

Eczema. A superficial inflammatory process involving primarily the epidermis; characterized early by redness, itching, minute papules and vesicles, weeping of the skin, oozing, and crusting and later by scaling, lichenification, and often pigmentation.

ED_{50}. The dose of a particular substance that elicits an observable response in 50% of the test subjects.

Edema. The presence of abnormally large amounts of fluid in intercellular spaces within a tissue.

Electrophile. A chemical compound or group that is attracted to electrons and tends to accept electrons.

Elimination. The removal of a chemical substance from the body. The rate of elimination depends on the nature of the chemical and the mechanisms that are used to remove the chemical from the organism. Examples of mechanisms include expiration from the lungs, excretion by the kidneys by way of the urinary system, excretion in the sweat or saliva, and chemical alteration by the organism and subsequent excretion by any of these mechanisms. See **Excretion.**

Emphysema. Literally, an inflation or puffing up; a condition of the lung characterized by an increase, beyond the normal, in the size of air spaces distal to the terminal bronchiolus.

Encephalopathy. Any degenerative disease of the brain.

Endoplasmic reticulum. An ultramicroscopic organelle of nearly all cells of higher plants and animals, consisting of a more or less continuous system of membrane-bound cavities that ramify throughout the cell cytoplasm.

Endothelial. Pertaining to the layer of flat cells lining blood and lymphatic vessels.

Endotoxin. A toxin produced by certain bacteria and released on destruction of the bacterial cell.

Engineering control. A method of controlling exposures to contaminants by modifying the source or reducing the quantity of contaminants released into the environment.

Enterohepatic circulation. The recurrent cycle in which the bile salts and other substances excreted by the liver pass through the intestinal mucosa and become reabsorbed by the hepatic cells and then are reexcreted.

Environmental toxicology. That branch of toxicology that deals with exposure of biologic tissue (more specifically, human life) to chemicals that are basically contaminants of the biologic environment, or of food, or of water. It is the study of the causes, conditions, effects, and limits of safe exposure to such chemicals.

Eosinophil. A structural cell or histologic element readily stained by eosin; especially, a granular leukocyte containing a nucleus usually with two lobes connected by a slender thread of chromatin and having cytoplasm containing coarse round granules that are uniform in size.

Epidemic. Spreading rapidly and extensively by infection and affecting many individuals in an area or a population at the same time: an epidemic outbreak of influenza.

Epidermal tumor. A tumor arising from the skin (dermal) epithelial layer.

Epidermis. The outermost and nonvascular layer of the skin. It derives from embryonic ectoderm.

Epistaxis. Nosebleed.

Epithelioma. Any tumor developing in the epithelium, which is the kind of tissue that covers internal and external surfaces of the body.

Epoxide. An organic compound containing a reactive group comprising a ring formed by an oxygen atom joined to two carbon atoms, having the structure at right.

Erethism. Excessive irritability or sensitivity to stimulation, particularly with reference to the sexual organs, but including any body parts. Also a psychic disturbance marked by irritability, emotional instability, depression, shyness, and fatigue, which are observed in chronic mercury poisoning.

Erythema. The redness of the skin produced by congestion of the capillaries.

Erythropoiesis. The production of erythrocytes (red blood cells).

Erythropoietic stimulating factor (ESF). A factor or substance that stimulates the production of erythrocytes; may be the same as erythropoietin.

Erythropoietin. A protein that enhances erythropoiesis.

Ester. A compound formed from an alcohol and an acid by removal of water.

Esterase. Any of various enzymes that catalyze the hydrolysis of an ester.

Ether. A colorless, transparent, mobile, very volatile liquid, highly inflammable, and with a characteristic odor; many ethers are used by inhalation as general anesthetics; the usual anesthetic forms are diethyl ether or ethyl ether.

Excretion. The process whereby materials are removed from the body to the external environment. If a chemical is in solution as a gas at body temperature, it will appear in the air expired from the animal; if it is a nonvolatile substance, it may be eliminated by the kidney via the urinary system, or it may be chemically altered by the animal and then excreted by means of any of the mechanisms available to the animal, such as excretion in the urine, in the sweat, or in the saliva. See **Elimination**.

Fibrosis. The formation of excessive fibrous tissue, as in a reparative or reactive process.

Follicle-stimulating hormone. One of the gonadotropic hormones of the anterior pituitary, which stimulates the growth and maturation of graafian follicles in the ovary and stimulates spermatogenesis in the male.

Forensic toxicology. The medical aspects of the diagnosis and treatment of poisoning and the legal aspects of the relationships between exposure to and harmful effects of a chemical substance. It is concerned with both intentional and accidental exposures to chemicals.

Gastric lavage. The process of washing out the stomach with saline solution using a lavage tube to remove poisons taken orally.

Gastritis. Inflammation of the stomach.

Gene. The basic unit of inheritance, recognized through its variant alleles; a segment of DNA coding a designated function (or related functions).

Genotoxicity. A measure of the potency of adverse effect of a substance directly of DNA.

Genotype. The entire allelic composition of an individual (or genome), or of a certain gene or set of genes.

Germ cell. An ovum or a sperm cell of one of its developmental precursors.

Gingivitis. Inflammation of the gums of the mouth.

Glutathione. A naturally occurring tripeptide, serving as a biological redox agent or substrate for certain conjugation reactions in chemical metabolism.

Glycoside. Any of a group of organic compounds, occurring abundantly in plants, that yield a sugar and one or more nonsugar substances on hydrolysis.

Gonadotropin. A hormone that stimulates the growth and activity of the gonads, especially any of several pituitary hormones that stimulate the function of the ovaries and testes.

Granulocyte. Any cell containing granules, especially a leukocyte containing neutrophil, basophil, or eosinophil granules in its cytoplasm.

Halogenation. The incorporation of one of the halogen elements, usually chlorine or bromine, into a chemical compound.

Heinz body. Microscopic bodies noted in red blood cells with enzyme deficiencies, identified as either cholesterin-olein-based or as dead cytoplasm resulting from oxidative injury to and precipitation of hemoglobin.

Hematopoietic. Pertaining to or affecting the formation of blood cells; an agent that promotes the formation of blood cells.

Hematuria. Blood in the urine.

Hemolytic. Pertaining to, characterized by, or producing hemolysis. The liberation of hemoglobin; the separation of hemoglobin from the red cells and its appearance in the plasma.

Hemolytic anemia. Anemia owing to shortened *in vivo* survival of mature red blood cells and inability of the bone marrow to compensate for their decreased lifespan.

Hemoptysis. The coughing or spitting up of blood from the respiratory tract.

Hemorrhagic cystitis. Urinary bladder inflammation compounded with bleeding.

Hemosiderosis. A general increase in iron stores in tissues without tissue damage.

Hepatomegaly. Enlargement of the liver.

Hepatotoxin. A toxin destructive of liver cells.

Histamine. A physiologically active amine, $C_5H_9N_3$, found in plant and animal tissue. It is released from cells of the immune system in human beings as part of an allergic reaction.

Hives. A skin condition characterized by intensely itching welts and caused by an allergic reaction to internal or external agents, an infection, or a nervous condition. Also called *urticaria.*

Homolog. One of a series of compounds, each of which is formed from the one before it by the addition of a constant element or a constant group of elements, as in the homologous series C_NH_{2N+2}, compounds of which would be CH_4, C_2H_6, C_3H_8, or similar.

Humoral immunity. The component of the immune response involving the transformation of B lymphocytes into plasma cells that produce and secrete antibodies to a specific antigen.

Hydrocarbon. An organic compound consisting exclusively of the elements carbon and hydrogen. Derived principally from vegetable sources, petroleum, and coal tar.

Hydrolysis. Decomposition of a chemical compound by reaction with water, such as the catalytic conversion of starch to glucose.

Hydrophilic. Readily absorbing water; hygroscopic.

Hydroxylation. An oxidative reaction that introduces one or more hydroxyl groups into an organic compound.

Hyperalgesia. A heightened or excessive sensitivity to pain.

Hyperemia. An excess of blood in some part of the body.

Hyperkeratosis. Overgrowth of the corneous layer of the skin or any disease characterized by that conditions.

Hyperpigmentation. Abnormally increased pigmentation.

Hyperplasia. Abnormal multiplication or increase in the number of normal cells in normal arrangement in a tissue.

Hypersensitivity. A state of extreme sensitivity to an action of a chemical; for example, the individuals of a test population who fit into the "low end" of an ED_{50} or LD_{50} curve (i.e., those individuals who react to a very low dose as opposed to the median effective dose).

Hypokinesis. Abnormally decreased mobility; abnormally decreased motor function or activity.

Hyposensitivity. The state of decreased sensitivity; for example, the individuals of a test population who fit into the "high end" of an ED_{50} or LD_{50} curve (i.e., those individuals who respond only to a very high dose as compared to the median effective dose).

Hypoxia. A partial reduction in the oxygen concentration supplied to cells or tissues.

Immune response. See **Sensitization reaction.**

Incidence. An expression of the rate at which a certain event occurs, as the number of new cases of a specific disease occurring during a certain period.

Inclusion body. An abnormal structure in a cell nucleus or cytoplasm having characteristics staining properties and associated especially with certain viral infections, such as rabies and smallpox.

Infarct. An area of necrosis in a tissue caused by local lack of blood resulting from obstruction of circulation to the area.

Inhalation route. The movement of a chemical from the breathing zone, through the air passageways of the lung, into the alveolar area, across the epithelial cell layer of the alveoli and the endothelial cell layer of the capillary wall, and into the blood system.

Inotropic. Affecting the force of muscular contraction, especially in the heart muscle.

Interleukins A generic term for a group of protein factors that affect primary cells and are derived from macrophages and T cells that have been stimulated by antigens or mitogens.

Interleukin-1. Any of a group of protein substances, released by macrophages and other cells, that induce the production of interleukin-2 by helper T cells and stimulate the inflammatory response.

Interleukin-2. A lymphokine that is released by helper T cells in response to an antigen and interleukin-1 and stimulates the proliferation of helper T cells. It has been used experimentally to treat cancer.

Intraperitoneal. Within the peritoneal cavity; an intraperitoneal injection is one in which a chemical is injected into the abdominal fluid of an animal.

Ionization. The dissociation of a substance in solution into ions.

Irritant dermatitis. See **Dermatitis**.

Ischemia. Deficiency of blood owing to a functional constriction or actual obstruction of a blood vessel.

Isotonic. Describing a solution with the same solute concentration as another solution (e.g., tissue culture media and cellular cytoplasm of cultured cells).

Isozyme. A member of a family of proteins with related structure and function.

Kepone. Insecticide and fungicide having the formula $C_{10}Cl_{10}O$; causes excitability, tremor, skin rash, opsoclonus, weight loss, and in some cases (in animals) testicular atrophy.

Keratoacanthoma. A rapidly growing papular lesion, with a crater filled with a keratin plug, which reaches maximum size and then resolves spontaneously within 4–6 months from onset.

Keratosis. Any horny growth, such as a wart or callosity.

Ketone. Any compound containing the carbonyl group C=O and having hydrocarbon groups attached to its carbonyl carbon.

LD_{50}. That dose of a particular substance that, administered to all animals in a test, is lethal to 50% of the animals. It is that dose of a compound that will produce death in 50% of the animals—hence, the median lethal dose. The values of LD_{50} should be reported in terms of the duration over which the animals were observed. If a time is not given, it is assumed they were observed for 24 h.

Lacrimation. The secretion and discharge of tears.

Laryngitis. Inflammation of the larynx, a condition attended with dryness and soreness of the throat, hoarseness, cough, and dysphagia (difficulty in swallowing).

Leukocyte. A white blood cell or corpuscle; classified as either granular or nongranular.

Leukocytosis. A transient increase in the number of leukocytes in the blood, resulting from various causes, such as hemorrhage, fever, infection, or inflammation.

Leukopenia. Lower-than-normal number of leukocytes in the blood; the normal concentration is 4,000–11,000 leukocytes in 1 ml of blood.

Leydig cells. The interstitial cells of the testes (between the seminiferous tubules), believed to furnish the male sex hormone.

Lichen planus. An inflammatory skin disease characterized by the appearance of wide, flat, violaceous, itchy, polygonal papules, occurring in circumscribed patches, and often very persistent. The hair follicles and nails may become involved, and the buccal mucosa may be affected.

Lipid peroxidation. Interaction of free radicals with the lipid constituents of a membrane, resulting in alterations of structure and function of the membrane.

Lipophilicity. Having an affinity for fats.

Lipoprotein. Any of a group of conjugated proteins in which at least one component is a lipid. Lipoproteins, classified according to their densities and chemical

qualities, are the principal means by which lipids are transported in the blood.

Locus of action (site of action). The part of the body (organ, tissue, or cell) where a chemical acts to initiate the chain of events leading to a particular effect.

Luteinizing hormone. A gonadotropic hormone of the anterior pituitary, which acts with the follicle-stimulating hormone to cause ovulation of mature follicles and secretion of estrogen by thecal and granulosa cells.

Lymphocyte. A mononuclear leukocyte with a deep-staining nucleus containing dense chromatin and a pale-blue-staining cytoplasm. Chiefly a product of the lymphoid tissue. Participates in humoral and cell-mediated immunity. See also **B cell and T cell**.

Lymphokine. Any of various substances released by T cells that have been activated by antigens. They function in the immune response through a variety of actions, including stimulating the production of nonsensitized lymphocytes and activating macrophages.

Macrophage. Any of the large phagocytic cells of the reticuloendothelial system.

Makeup air. In workplace ventilation, air introduced into an area to replace the air that has been removed.

Malignant tumor. Relatively autonomous growth of cells or tissue. Each type of malignant tumor has a different etiology and arises from a different origin. The condition tends to become progressively worse and to result ultimately in death. There are many common properties of malignant tumors, but the invasion of surrounding tissue and the ability to metastasize are considered the most characteristic.

Margin of safety. The magnitude of the range of doses involved in progressing from a noneffective dose to a lethal dose. Consequently, the slope of the dose–response curve is an index of the margin of safety of a compound.

Megakaryocyte. A giant cell found in **the** bone marrow, containing a greatly lobulated nucleus from which mature blood platelets originate.

Mesenchymal cells (tissue). The meshwork of embryonic corrective cells or tissue in the mesoderm from which are formed the connective tissues of the body, the blood vessels, and the lymphatic vessels.

Mesothelioma. A tumor developed from the mesothelial tissue—the simple squamous-celled layer of the epithelium, which covers the surface of all true serous membranes (lining the abdominal cavity, covering the heart, and enveloping the lungs).

Metabolism. The biochemical reactions that take place within an organism. It involves two processes: anabolism (assimilation or constructive processes) and catabolism (disintegration or destructive processes). All metabolic processes involve energy transfer.

Metallothionein. An inducible metal-binding protein involved in trafficking and detoxification mechanisms for various heavy metals.

Metaplasia. The transformation of cells from a normal to an abnormal state.

Metastasis. The establishment of a secondary growth site, distant from the primary site. One of the primary characteristics of a malignant tumor.

Methemoglobin. A compound formed from hemoglobin by oxidation of iron in the ferrous state to the ferric state. Methemoglobin does not combine with oxygen.

Methemoglobinemia. Presence of methemoglobin in the blood, resulting in cyanosis.

Microsomes. The fragments of the smooth reticular endothelium. This is the source of the microsomal enzymes that are capable of catalyzing a variety of biotransformation reactions, including hydroxylation, dealkylation, deamination, alkyl side-chain oxidation, hydrolysis, and reduction.

Miosis. Contraction of the pupil of the eye.

Mitochondria. Small spherical or rod-shaped components (organelles) found in the cytoplasm of cells, enclosed in a double membrane. They are the principal sites of energy generation (ATP), and they contain the enzymes of the Krebs and fatty acid cycles and the respiratory pathways. Mitochondria contain an extranuclear source of DNA and have genetic continuity.

Mitosis. The process in cell division by which the nucleus divides, typically consisting of four stages—prophase, metaphase, anaphase, and telophase—followed by cytokinesis and normally resulting in two new cells, each of which contains a complete copy of the parental chromosomes.

Mixed-function oxidase system (MFO). A nonspecific, multienzyme complex on the smooth endoplasmic reticulum of cells in the liver and various other tissues. These enzymes constitute the important enzyme system involved in phase I reactions (i.e., oxidation/reduction reactions). Also called *cytochrome P450 enzymes.*

Monoclonal antibody. A homogeneous antibody that is produced by a clone of antibody-forming cells and that binds with a single antigenic determinant.

Monocyte. A mononuclear phagocytic leukocyte with an ovoid or kidney-shaped nucleus, containing lacy, linear chromatin, and abundant gray-blue cytoplasm fitted with fine, reddish, and azure granules.

Morbidity. The rate of sickness or ratio of sick persons to well persons in community.

Multiple myeloma. A malignant proliferation of plasma cells in bone marrow causing numerous tumors and characterized by the presence of abnormal proteins in the blood.

Muscarine. A highly toxic alkaloid $C_9H_{20}NO_2$, related to the cholines, derived from the red form of the mushroom *Amanita muscaria.*

Mutagen. Any substance causing genetic mutation.

Mutagenesis. The induction of those alterations in the information content (DNA) of an organism or cell that are not due to the normal process of recombination. Mutagenesis

is irreversible and is cumulative, in the event of increased mutation rates or decreased selection pressures.

Mutagenic tests. Test of an agent to determine effects on the faithful replication of genetic material. The genetic damage can occur in both somatic and germinal cell lines.

Mutation. A permanent offspring-transmissible change in genetic material or structure. Such changes may manifest themselves as altered morphology or altered ability to direct the synthesis of proteins.

Myalgia. Muscular pain or tenderness, especially when diffuse and nonspecific.

Myasthenia gravis. A disease characterized by progressive fatigue and generalized weakness of the skeletal muscles, especially those of the face, neck, arms, and legs, caused by impaired transmission of nerve impulses following an autoimmune attack on acetylcholine receptors.

Myelin sheath. The insulating envelope of myelin that surrounds the core of a nerve fiber or axon and facilitates the transmission of nerve impulses. In the peripheral nervous system, the sheath is formed from the cell membrane of the Schwann cell and, in the central nervous system, from oligodendrocytes. Also called *medullary sheath.*

Myeloid leukemia. Leukemia arising from myeloid tissue (bone marrow) characterized by unrestrained growth of the granular, polymorphonuclear leukocytes and their precursors.

Myelotoxin. A cytotoxin that causes destruction of bone marrow cells.

Myoclonus. A sudden twitching of muscles or parts of muscles, without any rhythm or pattern, occurring in various brain disorders.

Myotonia. Tonic spasm or temporary rigidity of one or more muscles, often characteristic of various muscular disorders.

Narcosis. A condition of deep stupor or unconsciousness produced by a drug or other chemical substances.

Nasopharyngeal region. The part of the pharynx lying above the level of the soft palate (also known as the *postnasal space*).

Necrosis. Death of one or more cells, or of part of a tissue or organ, generally owing to irreversible damage.

Nematocyst. A capsule within specialized cells of certain coelenterates, such as jellyfish, containing a barbed, thread-like tube that delivers a paralyzing sting when propelled into attackers and prey.

Neoplasm. Literally, new growth, usually characterized by a random abnormal "immature," meiosis-type cell division and proliferation.

Nephritis. Inflammation of the kidney; a focal or diffuse proliferative or destructive process, which may involve the glomerulus, tubule, or interstitial renal tissue.

Neurodermatitis. A nonspecific pruritic skin disorder presumed to result from prolonged vigorous scratching, rubbing, or pinching, sometimes forming polymorphic lesions.

Neuroendocrine. Or, relating to, or involving the interaction between the nervous system and the hormones of the endocrine glands.

Neurofibril. One of the delicate threads running in every direction through the cytoplasm of the body of a nerve cell and extending into the axon and dendrites of the cell.

Neuromuscular endplate. A flattened discoid expansion at the neuromuscular junction, where a myelinated motor nerve fiber joins a skeletal muscle fiber.

Neuropathy. General term denoting functional disturbances and/or pathologic changes in the nervous system.

Neutropenia. A decrease in the number of neutrophilic leukocytes in the blood.

Neutrophil. A granular leukocyte having a nucleus with three to five lobes connected by slender threads of chromatin and cytoplasm, containing fine, inconspicuous granules.

Nicotinic effect. Poisoning by nicotine or a compound related in structure or action, characterized by stimulation (low doses) and depression (high doses) of the central and autonomic nervous systems. In extreme cases, death results from respiratory paralysis. Also referred to as *nicotinism.*

Nitrosamine. Any of a group of *n*-nitroso derivatives of secondary amines. Some show carcinogenic activity.

NOEL. See **No-observable-effect level.**

Noncompetitive inhibition. Inhibition of enzyme activity by inhibitors that combine with the enzyme on a site other than that utilized by the substrate; such inhibition may be irreversible or reversible.

Nonspecific chemical action. The action of a chemical, such as a strong acid or base or concentrated solution of organic solvent, which occurs in all cells in direct proportion to the concentration in contact with the tissue. This is a nonselective effect, and its intensity is directly related to the concentration of the chemical.

Nonspecific receptor. Secondary receptor within the body, which combine with or react with a chemical; however, the function of the cell is not influenced by the product that is formed. Such receptors are usually combining sites on proteins.

No-observable-effect level (NOEL). A measure of the toxicity of a substance, established by the U.S. Environmental Protection Agency (USEPA); the level of a substance that, when administered to a group of experimental animals, does not produce those effects observed at higher levels and at which no significant differences between the exposed animals and the unexposed or control animals are observed.

Olefin. A class of unsaturated aliphatic hydrocarbons having one or more double bonds. Also called *alkene.*

Oncogenic. Giving rise to tumors or causing tumor formation.

Opsoclonus. A condition characterized by rapid, irregular, nonrhythmic horizontal and vertical oscillations of the eyes, observed in various disorders of the brainstem or cerebellum.

Optic neuritis. Inflammation of the optic nerve; it may affect the part of the nerve within the eyeball, or the portion behind the eyeball.

Oral route. The entry of a chemical into the body by way of the gastrointestinal tract. Although absorption to some extent takes place throughout the tract, the majority of the absorption takes place in the area of the villi of the small intestine.

Organic acid. Any acid, the radical of which is a carbon derivative; a compound in which a hydrocarbon radical is joined to COOH (carboxylic acid) or to SO_3H (sulfonic acid).

Organochlorine pesticides. These compounds are extremely stable and persistent in the environment. They are efficiently absorbed by ingestion and act on the central nervous system to stimulate or depress it. Signs and symptoms of toxicity vary with the specific chemical. In general, mild poisoning cases cause symptoms such as dizziness, nausea, abdominal pain, and vomiting. In chronic poisoning, weight loss and loss of appetite, temporary deafness, and disorientation can occur.

Organophosphate pesticides. These are irreversible inhibitors of cholinesterase, thus allowing accumulating of acetylcholine at nerve endings. They are rapidly absorbed into the body by ingestion, through intact skin, including the eye, and by inhalation. Poisoning symptoms range from headache, fatigue, dizziness, vomiting, and cramps in mild cases to the rapid onset of unconsciousness, local or generalized seizure, and other manifestations of a cholinergic crisis in severe cases.

Osteomalacia. A condition of softening of the bones characterized by pain, tenderness, loss of weight, and muscular weakness.

Osteoporosis. Abnormal rarefaction of the bone, seen most commonly in the elderly.

Osteosclerosis. Hardening or abnormal density of bone.

Ototoxic. Having a toxic effect on the structures of the ear, especially on its nerve supply.

Pancytopenia. A form of anemia in which the capacity of the bone marrow to generate red blood cells is defective. This anemia may be caused by bone marrow disease or exposure to toxic agents, such as radiation, chemicals, or drugs.

Paraffin. A class of aliphatic hydrocarbons characterized by a straight or branched carbon chain afid having the generic formula C_nH_{2n+2}; also called *alkane.*

Paranoid schizophrenia. A psychotic state characterized by delusions of grandeur or persecution, often accompanied by hallucinations.

Parasympathetic. Craniosacral division of the autonomic nervous system. These cholinergic nerves are associated with normal body functions (e.g., smooth muscle in blood vessels, salivary glands, and GI tract).

Paresthesia (also paraesthesia). A skin sensation, such as burning, prickling, itching, or tingling, with no apparent physical cause.

Parkinsonism. A group of neurologic disorders characterized by hypokinesia, tremor, and muscular rigidity.

PEL. See **Permissible exposure limit.**

Percutaneous absorption. The transfer of a chemical from the outer surface of the skin through the horny layer (dead cells), through the epidermis, and into the systemic circulation. A variety of factors, such as pH, extent of ionization, molecular size, and water and lipid solubility govern transfer of chemicals through the skin.

Perinatal toxicology. The study of toxic responses to occupationally or environmentally encountered substances during a woman's exposure to those substances from the time of conception through the neonatal period.

Peripheral neuritis. Inflammation of the nerve ending or of terminal nerves.

Permissible exposure limit (PEL). A measure of the toxicity of a substance, established by the U.S. Department of Labor, Occupational Safety and Health Administration (OSHA); an 8 h time-weighted average (TWA) limit of exposure is most commonly used. The limit is commonly expressed as the concentration of a substance per unit of air volume (mg/m^3, ppm, $fibers/cm^3$, etc.).

Pernicious anemia. The progressive, megaloblastic anemia resulting from lack of vitamin B_{12}, sometimes accompanied by degeneration of the posterior and lateral columns of the spinal cord.

Peroxidase. Any of a group of enzymes that occur especially in plant cells and catalyze the oxidation of a substance by a peroxide.

Personal protective equipment. Any devices worn by individuals as protection against hazards in the environment or the workplace, including respirators, gloves, goggles, and earmuffs.

Pesticide. Any substance used to destroy or inhibit the action of plant or animal pests. See **Carbamate**, **Organochlorine pesticides**, and **Organophosphate pesticides**.

Petechiae. Tiny, nonraised, perfectly round, purplish red spots caused by intradermal or submucosal hemorrhaging.

pH. A value taken to represent the acidity or alkalinity of an aqueous solution. It is defined as the logarithm of the reciprocal of the hydrogen ion concentration of a solution:

$$\mathrm{pH} = \ln \frac{1}{\left[\mathrm{H}^+\right]}$$

Pharmacokinetics. The field of study concerned with the techniques used to quantify the absorption, distribution, metabolism, and excretion of drugs or chemicals in animals, as a function of time.

Pharmacology. The unified study of the properties of chemical agents (drugs) and living organisms and all

aspects of their interactions. An expansive science encompassing areas of interest germane to many other disciplines.

Phenothiazine. A green, tasteless compound with the formula $C_{12}H_9NS$, prepared by fusing diphenylamine with sulfur; also, a group of tranquilizers resembling phenothiazine in molecular structure.

Phocomelia. A developmental anomaly characterized by the absence of the proximal portion of a limb or limbs, such that hands or feet are attached to the trunk of the body by a single small, irregularly shaped bone.

Phospholipase. Any enzyme that catalyzes the hydrolysis of a phospholipid.

Photophobia. Abnormal sensitivity, usually of the eyes, to light.

Photosensitivity reactions. Undesirable reactions in the skin of persons exposed to certain chemicals when the skin is also exposed to sunlight (in some cases, to artificial light). Dermatologic lesions form, which vary from sunburn-like responses to edematous, vesiculated lesions or bullae.

Phototoxicity. Capacity of a chemical to (nonimmunologically) sensitize the skin to a light-induced reaction.

Pilosebaceous units. Relating to the hair follicles and sebaceous glands.

pK. The acidic dissociation constant of a compound; the pH of an aqueous solution of an acid or base at which equal concentrations of each are present, at the point at which dissociation is half-complete. The negative logarithm of the ionization constant K_a.

Pneumoconiosis. Accumulation of dusts in the lungs and the tissue reaction to the presence of such dust.

Pneumonitis. Inflammation of lung tissue.

Point mutation. An alteration in a single nucleotide pair in the DNA molecule, usually leading to a change in only one biochemical function.

Poison. The term used to describe those materials or chemicals that are distinctly harmful to the body.

Polymorphism. The occurrence of different genetic forms or types of a protein that produce phenotypically distinct populations.

Polymorphonuclear. Having a nucleus deeply lobed or so divided as to appear to be multiple.

Polyneuritis. Inflammation of several nerves simultaneously, as in lead palsy.

Polyneuropathy. A disease involving several nerves.

Porphyrin. Any of a group of iron-free or magnesium-free cyclic tetrapyrrole derivatives occurring universally in the protoplasm. They form the basis of the respiratory pigments of animals and plants.

Potency. A comparative expression of chemical or drug activity measured in terms of the dose required to produce a particular effect of given intensity relative to a given or implied standard of reference. If two chemicals are not both capable of producing an effect of equal magnitude, they cannot be compared with respect to potency.

Potentiation. A condition whereby one substance is made more potent in the presence of another chemical that alone produces no response.

Pressure, static. The potential pressure exerted in all directions by a fluid at rest.

Pressure total. The algebraic sum of static and velocity pressures, representing the total energy in the system.

Pressure, velocity. The kinetic pressure exerted in the direction of flow necessary to cause a fluid at rest to flow at a given velocity.

Primary carcinogens. Chemicals that act directly and without biotransformation. Also called *direct carcinogens.*

Primary irritants. Chemicals that induce local, minor to severe inflammatory response, or even extreme necrosis, of cells of a tissue, in direct relation to the concentration available to the tissue. This is termed a *nonspecific chemical action*, the toxicity of which may be manifested at the site of exposure (e.g., skin or in the respiratory tract). Examples of these types of chemicals are strong acids or bases, ammonia, and acrolein.

Probenecid. A white, odorless crystalline powder, with the formula $C_{13}H_{19}NO_4S$, soluble in dilute alkali, alcohol, and acetone; used to increase serum concentrations of certain antibiotics, as well as being an agent to promote uric acid secretion in the urine.

Procarcinogen. Chemicals that require metabolism to another more reactive or toxic chemical form before their carcinogenic action can be expressed.

Proerythropoietin. A precursor of erythropoietin.

Psoriasis. A chronic, hereditary, recurrent, papulosquamous dermatitis, the distinctive lesion of which is a vivid red macula, papule, or plaque covered almost to its edge by silvery lamellated scales. It usually involves the scalp and extensor surfaces of the limbs, especially the elbows, knees, and shins.

Pyrethroid. Any of several synthetic insecticidal compounds similar to the nature pyrethrums extracted from crushed chrysanthemums.

Raynaud's phenomenon. Intermittent attacks of severe pallor of the fingers or toes and sometimes of the ears and nose, brought on characteristically by cold and sometimes by emotion.

Receptors. See **Specific receptor** and **Nonspecific receptor**.

Renal osteodystrophy. A condition resulting from chronic kidney disease. The onset early in childhood is characterized by impaired renal function, elevated serum phosphorus and low or normal serum calcium levels, and stimulation of parathyroid function. The resultant bone disease includes a variety of symptoms, including osteitis fibrosa cystica, osteomalacia, osteoporosis, and

osteosclerosis. Renal dwarfism may result from childhood onset.

Reproduction tests. Tests that determine (or estimate) the effects of an agent on fertility, gestation, and offspring; usually conducted on more than one generation of test animals. Toxicity in either parent may affect fertility as the direct result of altered gonadal function, estrus cycle, mating behavior, and conception rates. Effects on gestation concern the development of the fetus. Effects on offspring concern growth, development, and sexual maturation; and effects on the mother concern lactation and acceptance of the offspring.

Resorption. The loss of substance in the mucous lining of the uterus.

Reticuloendothelial system. Phagocytic macrophages present in linings of sinuses and in the reticulum of various organs and tissues. A functionally important body defense mechanism; the phagocytic cells have both endothelial and reticular attributes and the ability to take up particles of colloidal dyes.

Risk assessment. A methodologic approach in which the toxicities of a chemical are identified, characterized, and analyzed for dose–response relationships, and a mathematical model is applied to the data to generate a numerical estimate that can serve as a guide to allowable exposures.

Risk estimation. Mathematical modeling of the animal and/or human toxicity data, combined with evaluation of human exposures, so as to estimate the probability or incidence of effects on human health.

Risk management. The process of applying a risk assessment to the conditions that exist in society, so as to balance exposures to toxic agents against needs for products and processes that may be inherently hazardous.

Safety factor. A factor that presumably reflects the uncertainties inherent in the process of extrapolating data about toxic exposures (i.e., intraspecies and interspecies variations). With this approach, an allowable human exposure to a compound can be determined by dividing the no-observable-effect level (NOEL) established in chronic animal toxicity studies by some safety factor. Also called *uncertainty factor.*

Sarcoma. A cancer that arises from mesodermal tissue (supporting or connective tissue).

Sclerodermatous skin change. A chronic hardening and shrinking of the connective tissues of any part of the body, including the skin, heart, esophagus, kidney, and lung. The skin may become thickened and hard, and the condition may be generalized and rigid, and pigmented patches may occur, limited to the distal parts of the extremities and face, or to the digits, or localized to oval or linear areas a few centimeters in diameter.

Sclerosis. A thickening or hardening of a body part, as of an artery, especially from excessive formation of fibrous interstitial tissue.

Sensitization reaction. An immunologic response to a chemical. The mechanism of immunization involves the following events: initial exposure of an animal to a chemical substance, an induction period in the animal; and the production of a new protein termed an *antibody.* The initial exposure does not result in cellular damage but causes the animal to be "sensitized" to subsequent exposure to the chemical. Exposure of the animal to the same chemical on a subsequent occasion will lead to the formation of sensitized antigen, which will react with the preformed antibodies and lead to a response in the tissues in the form of cellular damage.

SGOT. Serum glutamic oxaloacetic transaminase. An enzyme found in the liver and muscle tissue and used to detect early membrane permeability as part of a test of the activity of enzymes present in liver cells.

SGPT. Serum glutamic pyruvic transaminase. An enzyme used in the identification and measurement of the activity of enzymes present in liver cells. SGPT is found in the liver and heart tissues. It is an indicator of early membrane permeability, as is SGOT.

Silicosis. A type of pneumoconiosis due to the inhalation of the dust of stone, sand, or flint containing silicon dioxide. It results in the formation of generalized nodular fibrotic changes in both lungs.

Site of action. See **Locus of action.**

SNARL. Suggested no-adverse-response level. A measure of toxicity established by the National Research Council.

Somatic cell. Any cell of a plant or an animal other than a germ cell. Also called *body cell.*

Specific receptor. Macromolecular constituent of tissue capable of combining reversibly with a compound by means of chemical bonds; the tissue element with which a compound interacts to provide its characteristic biologic effect.

Spermatozoa. The mature fertilizing gamete of a male organism, usually consisting of a round or cylindrical nucleated cell, a short neck, and a thin, motile tail. Also called *sperm cell.*

Spirometer. An instrument for measuring the air taken into and exhaled from the lungs.

Spirometry. The measurement of the breathing capacity of the lungs.

Squamous cell carcinoma. Carcinoma developing from squamous epithelium (composed of flattened, platelike cells) and characterized by cuboid cells.

Stereoisomers. Two substances of the same composition differing only in the relative spatial positions of their constituent atoms and/or groups.

Steric hindrance. The nonoccurrence of an expected chemical reaction owing to inhibition by a particular atomic grouping.

Sulfhemoglobin. An abnormal greenish form of hemoglobin containing sulfur that is bound to heme.

Sulfotransferase. An enzyme that catalyzes the transfer of sulfate from a donor molecule to an acceptor.

Sympathetic. Thoracolumbar division of the autonomic nervous system mainly involved with homeostasis (e.g., vasoconstriction, glucose mobilization, adrenaline release).

Sympathomimetic. Mimicking the effects of impulses conveyed by adrenergic postganglionic fibers in the sympathetic nervous system.

Synapse. The anatomical relation of one nerve cell to another; the region of junction between processes of two adjacent neurons, forming the place where a nervous impulse is transmitted from one neuron to another.

Synergism. The situation in which the combined effects on a biologic system of two chemicals acting simultaneously are greater than the algebraic sum of the individual effects of these chemicals.

Tachycardia. A rapid heart rate, especially one above 100 beats per minute in an adult.

Tachypnea. Rapid breathing.

T cell. Thymus-dependent lymphocytes; these pass through the thymus or are influenced by it on their way to the tissues; they can be suppliers or assist the stimulation of antibody production in B cells in the presence of antigen and can kill such cells as tumor and transplant tissue cells. T cells are responsible for all cell-mediated immunity and immunologic memory.

TD_{50}. That dose of a substance that, administered to all animals in a test, produces a toxic response in 50% of them. The toxic response may be any adverse effect other than death.

Teratogen. Any substance capable of causing malformation during development of the fetus.

Thalidomide. A sedative and hypnotic drug commonly used in Europe in the early 1960s. It was discovered to be the cause of serious congenital anomalies in the fetus, notably amelia and phocomelia.

Threshold dose (ThD). The minimal dose effective in prompting an all-or-none response.

Threshold limit value (TLV). A term for exposure limits established by the American Conference of Governmental Industrial Hygienists (ACGIH). That concentration of any airborne substance to which it is believed, through animal toxicity testing and human exposure data, that workers can be exposed to 8 h per day, 40 h per week for a working lifetime, without suffering adverse health effects or significant discomfort. TLV measurements are usually based on 8 h time-weighted average (TWA) exposures but may be expressed as ceiling values.

Time-weighted average (TWA). A method of combining multiple air sample results collected on one individual during a work shift, so as to derive the overall average exposure for the entire shift (or exposure period). Measurements of chemical exposure can be made in each phase, and the exposure estimate is calculated according to the formula

$$E = \frac{C_1T_1 + C_2T_2 + \cdots + C_NT_N}{T_1 + T_2 + \cdots + T_N}$$

where is the E exposure, C is the concentration measured in phase N, and T_N = duration of phase N.

TLV. See T**hreshold limit value**.

Tolerance. The ability of an organism to show less response to a specific dose of a chemical than it showed on a prior occasion when subjected to the same dose.

Totipotency. The ability of a cell, such as an egg, to give rise to unlike cells and thus to develop into or generate a new organism or part.

Toxicity. A relative term generally used in comparing the harmful effect of one chemical on some biologic mechanism with the effect of another chemical.

Toxicology. The scientific study of poisons and their actions and detection and treatment of the condition produced by them. Also the study of the effects of chemicals on biologic systems, with emphasis on the mechanisms of harmful effects of chemicals and the conditions under which harmful effects occur. Thus, toxicology is a multidisciplinary science.

Tracheitis. Inflammation of the trachea.

Tumor. An abnormal mass of tissue, the growth of which exceeds and is uncoordinated with that of normal tissue. The basic types are benign and malignant. See also **Benign tumor** and **Malignant tumor**.

Tumorigenicity. The process or quality of producing or giving rise to a tumor.

Type I reaction (immediate). A hypersensitive reaction that is manifested in allergic asthma, hay fever, and eczema, developing within minutes after exposure to an antigen.

Type ll reaction (antibody mediated). A hypersensitive reaction that is caused by an antibody reaction to cell surface antigens; occurs when red blood cells are destroyed in transfusion reactions.

Ubiquinone. A quinone compound that serves as a electron carrier between flavoproteins and in cellular respiration.

Uncertainty factor. See **Safety factor**.

Unction. An ointment; the application of an ointment or salve.

Uropathy. Any pathologic change in the urinary tract.

Urticaria. See **Hives**.

Ventricular fibrillation. Arrhythmia characterized by fibrillary contractions of the ventricular muscle owing to rapid repetitive excitation of myocardial fibers without coordinated contraction of the ventricle.

Xenobiotic. A chemical foreign to the biologic system (i.e., chemicals that are not normal endogenous compounds for the biologic system).

INDEX

Note: Page numbers in *italics* refer to Figures; those in **bold** to Tables

Principles of Toxicology: Environmental and Industrial Applications, Third Edition. Edited by Stephen M. Roberts, Robert C. James and Phillip L. Williams.

WILEY END USER LICENSE AGREEMENT

Go to www.wiley.com/go/eula to access Wiley’s ebook EULA.