ACADEMIC PRESS

TOXICOLOGY

Edited by

Hans Marquardt

Siegfried G. Schäfer

Roger O. McClellan

Frank Welsch

TOXICOLOGY

This Page Intentionally Left Blank

TOXICOLOGY

Edited by

Hans Marquardt
Department of Toxicology
Hamburg University Medical School and
Department of Toxicology and Environmental
Medicine of the Fraunhofer Society
Hamburg, Germany

Siegfried G. Schäfer
Solvay Pharmaceuticals
Hannover, Germany and
Department of Toxicology
Hamburg University Medical School
Hamburg, Germany

Roger McClellan
Chemical Industry Institute of Toxicology
Research Triangle Park
North Carolina

Frank Welsch
Chemical Industry Institute of Toxicology
Research Triangle Park
North Carolina

Academic Press
An Imprint of Elsevier

Front cover photograph: The monument of Paracelsus, father of toxicology, who lived worked in Salzburg, Austria and who also is buried there. Photograph taken by Dr. Herbert V. Levinsky. Background copyright Digital Stock Incorporated 1999.

This book is printed on acid-free paper. ♾

Academic Press
An Imprint of Elsevier
525 B Street, Suite 1900, San Diego, California 92101-4495, USA
http://www.apnet.com

Academic Press
24-28 Oval Road, London NW1 7DX, UK
http://www.hbuk.co.uk/ap/

Library of Congress Catalog Card Number: 98-89314

International Standard Book Number: 0-12-473270-4

PRINTED IN THE UNITED STATES OF AMERICA
04 MM 9 8 7 6 5 4 3

Contents

CHAPTER

23

Gastrointestinal Tract

KLAUS SCHÜMANN, BERND ELSENHANS, AND ELMAR RICHTER

CHAPTER

24

Cardiovascular System

PAUL J. BOOR

PART

III

TOXICOLOGY OF COMPOUNDS

CHAPTER

25

Hydrocarbons

GÜNTER KOSS AND IRENE TESSERAUX

CHAPTER

26

Halogenated Hydrocarbons

HERMANN M. BOLT AND JÜRGEN T. BORLAK

CHAPTER

27

Fluorocarbons

STEPHEN MAGDA

CHAPTER

28

Pesticides

RANDY L. ROSE, ERNEST HODGSON, AND R. MICHAEL ROE

This Page Intentionally Left Blank

Contributors

Numbers in parentheses indicate the pages on which the authors' contributions begin.

Bruce N. Ames (1) Division of Biochemistry and Molecular Biology, University of California at Berkeley, Berkeley, California 94720

Klaus Andreas (463) Institute of Pharmacology and Toxicology, Technical University Dresden, University Hospital, 01109 Dresden, Germany

Klaus E. Appel (1139) Federal Institute for Health Protection of Consumers and Veterinary Medicine, 14191 Berlin, Germany

Michael Arand (83) Institute of Toxicology, University of Mainz, 55131 Mainz, Germany

Steven I. Baskin (851) Division of Pharmacology, U.S. Army Medical Research Institute of Chemical Defense, Aberdeen Proving Ground, Maryland 21010

Hans Bigalke (1109) Medizinische Hochschule Hannover, Institut für Toxicologie, Zentrum Pharmakologie und Toxikologie, 30625 Hannover, Germany

Hermann M. Bolt (645) Institute of Occupational Physiology, University of Dortmund, 44139 Dortmund, Germany

Paul J. Boor (587) Department of Pathology, University of Texas Medical Branch, Galveston, Texas 77555

Jürgen T. Borlak (645) Fraunhofer Institute of Toxicology and Aerosol Research, 30625 Hannover, Germany

David J. Brusick (127) Covance Laboratories, Inc., Vienna, Virginia 22182

Richard J. Bull (1041) Pacific Northwest National Laboratory, Richland, Washington 99352

Wilfred Bursch (179) Institute of Cancer Research, University of Vienna, 1090 Vienna, Austria

Flemming R. Cassee (257) National Institute of Public Health and the Environment, Laboratory of Health Effects Research, 3720 BA Bilthoven, The Netherlands

Constantin Cojocel (297) Hoechst Marion Roussel, 65812 Bad Soden am Taunus, Germany

John F. Corbett (899) Bristol-Myers Squibb Worldwide Beauty Care, Scientific and Technical Affairs, Stamford, Connecticut 06922

Dorothee Charlotte Dartsch (245) Department of Toxicology and Environmental Medicine of the Fraunhofer Society, 20146 Hamburg, Germany

R. L. F. Dawes (755) Solvay Pharmaceuticals, 1381 CP Weesp, The Netherlands

Chris T. De Rosa (257) Division of Toxicology, Agency for Toxic Substances and Disease Registry, U.S. Department of Health and Human Services, Atlanta, Georgia 30333

William E. Dressler (899) Bristol-Myers Squibb Worldwide Beauty Care, Stamford, Connecticut 06922

Gerhard Eisenbrand (743) Department of Food Chemistry and Environmental Toxicology, University of Kaiserslautern, 67663 Kaiserslautern, Germany

Bernd Elsenhans (573, 755) Walther Straub Institute of Pharmacology and Toxicology, Ludwig Maximilians University, 80336 Munich, Germany

Peter Eyer (349, 805) Walther Straub Institute of Pharmacology and Toxicology, Ludwig Maximilians University, 80336 Munich, Germany

Victor J. Feron (257) TNO Nutrition and Food Research Institute, 3700 AJ Zeist, The Netherlands

Burckhard Fichtl (43) Walther Straub Institute of Pharmacology and Toxicology, Ludwig Maximilians University, 80336 Munich, Germany

Wolfgang Forth (755) Walther Straub Institute of Pharmacology and Toxicology, Ludwig Maximilians University, 80336 Munich, Germany

Heidi Foth (1067) Institute of Environmental Toxicology, Martin Luther University of Halle-Wittenberg, 06097 Halle, Germany

R. J. Michael Fry (937) Life Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37830

Heinz-Peter Gelbke (1116) Department of Toxicology, BASF, 67056 Ludwigshafen, Germany

Lois Swirsky Gold (1) Division of Biochemistry and Molecular Biology, University of California at Berkeley, Berkeley, California 94720

Alan M. Goldberg (1139) Center for Alternatives to Animal Testing, Johns Hopkins University, Baltimore, Maryland 21202

Gio B. Gori (1139) The Health Policy Center, Bethesda, Maryland 20816

Martin Göttlicher (231) Karlsruhe Research Center, Institute for Genetics, Karlsruhe, Germany

John P. Groten (257) TNO Nutrition and Food Research Institute, Toxicology Division, 3700 AJ Zeist, The Netherlands

Werner Grunow (1103) Bundesinstitut für gesundheitlichen Verbraucherschutz und Veterinärmedizin, 14195 Berlin, Germany

Ulrich F. W. Hagen (937) Karwinskistrasse 64, 81247 Munich, Germany

Gerald Hajian (25) Schering-Plough Research Institute, Kenilworth, New Jersey 07033

Georgia A. Hart (833) Health Safety and Environmental Department, Schuller International, Inc., Littleton, Colorado 80162

David G. Hattan (1087) Division of Health Effects Evaluation, Office of Premarket Approval, U.S. Food and Drug Administration Center for Food Safety and Applied Nutrition, Washington, DC 20204

Thomas W. Hesterberg (833) Health Safety and Environmental Department, Schuller International, Inc., Littleton, Colorado 80162

Bernhard Hildebrand (1116) Department of Toxicology, BASF, 67056 Ludwigshafen, Germany

Ernest Hodgson (663) Department of Toxicology, North Carolina State University, Raleigh, North Carolina 27695

Dietrich Hoffmann (898) Naylor Dana Institute for Disease Prevention, American Health Foundation, Valhalla, New York 10595

Ludwig A. Hothorn (25) University of Hannover, 30149 Hannover, Germany

Birgit Jödicke (491) 35581 Wetzlar-Münchholzhausen, Germany

Regine Kahl (273) Institute of Toxicology, University of Düsseldorf, 40225 Düsseldorf, Germany

Edward J. Kerfoot (1116) Product Regulations and Toxicology, BASF, Southgate, Michigan 48195

Reinhard Klimmek (349) Walther Straub Institute of Pharmacology and Toxicology, Ludwig Maximilians University, 80336 Munich, Germany

Werner Koehl (743) Department of Toxicology, Henkel, 40191 Düsseldorf, Germany

Günter Koss (603, 699) Department of Labor, Health and Social Affairs, 20148 Hamburg, Germany; and Department of Toxicology, Hamburg University Medical School, 20146 Hamburg, Germany

Johann Kummermehr (937) Institute for Radiobiology, National Center for Environment and Health, 85758 Neuherberg, Germany

Edward A. Lock (297) Central Toxicology Laboratory, Zeneca Ltd., Alderley Park, Macclesfield, Cheshire SK10 4TJ, United Kingdom

Eckhard Löser (919) Institute of Industrial Toxicology, Bayer, 42096 Wuppertal, Germany

Stephen Magda (657) Solvay Pharmaceuticals, 30002 Hannover, Germany

Brigitte Marian (179) Institute of Cancer Research, University of Vienna, 1090 Vienna, Austria

Hans Marquardt (151) Department of Toxicology, Hamburg University Medical School, 20146 Hamburg, Germany; and Department of Toxicology and Environmental Medicine of the Fraunhofer Society, 20146 Hamburg, Germany

Roger O. McClellan (331) Chemical Industry Institute of Toxicology, Research Triangle Park, North Carolina

Hans F. Merk (459) Department of Dermatology and Allergology, University Hospital, RWTH Aachen, 52074, Germany

Dietrich Milles (11) Center for Social Policy Research, University of Bremen, 28334 Bremen, Germany

Nancy A. Monteiro-Riviere (439) Center for Cutaneous Toxicology and Residue Pharmacology, North Carolina State University, Raleigh, North Carolina 27606

Hartwig Muhle (331) Fraunhofer Institute of Toxicology and Aerosol Research, 30625 Hannover, Germany

Diether Neubert (371, 491, 1153) Department of Toxicology, Institute of Clinical Pharmacology and Toxicology, Benjamin Franklin Medical Center, Free University of Berlin, 14195 Berlin, Germany

Reinhard Neubert (371) Groupement de Rechereche Scientifique et Médicale AHCRAA, Département Neurochirurgie, 74000 Annecy, France

Franz Oesch (83) Institute of Toxicology, University of Mainz, 55131 Mainz, Germany

Inga Ollroge (1201) Department of Labor, Health, and Social Affairs, 20148 Hamburg, Germany

Dennis James Paustenbach (1057) Exponent, Menlo Park, California

R. Julian Preston (937) Chemical Industry Institute of Toxicology, Research Triangle Park, North Carolina 27709

David E. Ray (463) Neurotox Section, MRC Toxicology Laboratory, University Leicester, Leicester LE13YD, England

Elmar Richter (573, 729) Walther Straub Institute of Pharmacology and Toxicology, Ludwig Maximilians University, 80336 Munich, Germany

Jim E. Riviere (439) Center for Cutaneous Toxicology and Residue Pharmacology, North Carolina State University, Raleigh, North Carolina 27606

Thomas E. Rodriguez (1051) McLaren–Hart, Philadelphia, Pennsylvania 19133

R. Michael Roe (663) Department of Entomology and Toxicology, North Carolina State University, Raleigh, North Carolina 27695

Randy L. Rose (663) Department of Toxicology, North Carolina State University, Raleigh, North Carolina 27695

Hugo W. Rüdiger (1027) Department of Occupational Medicine, University of Vienna, 1090 Vienna, Austria

Alan M. Rulis (1087) Office of Premarket Approval, U.S. Food and Drug Administration Center for Food Safety and Applied Nutrition, Washington, DC 20204

Gabriele Sabbioni (729) Walther Straub Institute of Pharmacology and Toxicology, Ludwig Maximilians University, 80336 Munich, Germany

Stephen H. Safe (1191) Department of Veterinary Physiology and Pharmacology, Texas A & M University, College Station, Texas 77843

Siegfried G. Schäfer (755) Solvay Pharmaceuticals, 30002 Hannover, Germany; and Department of Toxicology, Hamburg University Medical School, 20146 Hamburg, Germany

Rolf Schulte-Hermann (179) Institute of Cancer Research, University of Vienna, 1090 Vienna, Austria

Klaus Schümann (573, 755) Walther Straub Institute of Pharmacology and Toxicology, Ludwig Maximilians University, 80336 Munich, Germany

Doris A. Schwindt (459) Department of Dermatology and Allergology, University Hospital, RWTH 52074 Aachen, Germany

Raj K. Sharma (899) Bristol-Myers Squibb Worldwide Beauty Care, Scientific and Technical Affairs, Stamford, Connecticut 06922

Arpad Somogyi (1139) Federal Institute for Health Protection of Consumers and Veterinary Medicine, 14191 Berlin, Germany

Horst Spielmann (1139) National Center for Documentation and Evaluation of Alternatives to Animal Experiments, 12277 Berlin, Germany

Gisela Stropp (919) Institute of Industrial Toxicology, Bayer, 42096 Wuppertal, Germany

Ladislaus Szinicz (851) Institut für Pharmakologie und Toxikologie, Akademie des Sanitäts und Gesundheitswesens der Bundeswehr, 85748 Garching–Hochbrück, Germany

Raymond W. Tennant (217) Laboratory of Environmental Carcinogens, NIEHS, Research Triangle Park, North Carolina 27709

Irene Tesseraux (603) Department of Labor, Health, and Social Affairs, 20148 Hamburg, Germany

John A. Thomas (559) Department of Pharmacology, The University of Texas Health Science Center, San Antonio, Texas 78284

Michael J. Thomas (559) Department of Internal Medicine, Division of Endocrinology, University of North Carolina School of Medicine, Chapel Hill, North Carolina 27599

Kenneth R. Tindall (217) Laboratory for Environmental Carcinogens, NIEHS, Research Triangle Park, North Carolina 27709

Peter J. van Bladeren (251) TNO Nutrition and Food Research Institute, 3700 AJ Zeist, The Netherlands

David B. Warheit (833) Haskell Laboratory, E. I. du Pont de Nemours & Company, Newark, Delaware 19714

Hans-Herbert Wellhöner (1009) Medizinische Hochschule, Zentrum Pharmakologie und Toxikologie, 30625 Hannover, Germany

Frank Welsch (491) Chemical Industry Institute of Toxicology, Research Triangle Park, North Carolina 27709

Johannes Westendorf (959) Department of Toxicology, University Medical School, 20146 Hamburg, Germany

Detlef Wölfle (699) Federal Institute for Health Protection of Consumers and Veterinary Medicine, 14191 Berlin, Germany; and Department of Toxicology, Hamburg University Medical School, 20146 Hamburg, Germany

Ernst L. Wynder (898) Naylor Dana Institute for Disease Prevention, American Health Foundation, Valhalla, New York 10595

Maged Younes (111) International Programme on Chemical Safety, World Health Organization, 1211 Geneva 27, Switzerland

Preface

Toxicology, the science of poisons and poisonous activities, is an applied science. It describes chemical and biological interactions that have acute and chronic health-threatening effects, particularly on humans. It aims at quantifying these effects in order to detect, treat, and possibly prevent any damage. Thus, this discipline is challenged to make a significant scientific contribution to preventive medicine. ("The ultimate goal in medicine is to achieve to die young as late as possible," E. Wynder.) This was not always the case. The Greek words for bow and medicinal drug are "toxon" and "pharmakon"; hence, arrow poison is "toxon pharmakon." The historian Herodotus used the word "toxikon" for poison arrow. Combining this with "logos" for science gives the term "toxicology."

The study of poisons was an aspect of many ancient cultures (e.g., Egyptians—cardiac glycosides; Chinese—opium alkaloids; Incas—coca and strychnos alkaloids; Greeks—hemlock), and toxicology thus was descriptive and empirical. In contrast, modern toxicology, with an emphasis on mechanisms of action and toxicokinetics, is a fundamentally new science that is influencing medicine and greatly improving society. The 1990s particularly have seen a rapid expansion of knowledge regarding the biological basis of toxic reactions in response to exposure to xenobiotics. Molecular and cell biology have revealed a wealth of information on cellular mechanisms, such as signal pathways, as well as an increasing understanding of the consequences of loss of these controls. Emerging techniques, such as the use of transgenic animals, promise to stimulate continued progress. However, while toxicology has developed beyond describing reactions to compounds and is increasingly successful in unraveling the cellular and molecular bases of such reactions, the methods and practices of health risk assessment are mostly underdeveloped and have made little specific use of the mechanistic approaches to toxicology. Nevertheless, the ultimate goal of toxicological research is to provide a rational basis for judgments with regard to the potential health hazards of chemical exposure for humans and the environment. Because risk assessment is a process that has an impact on society at large, it must be based on a foundation as sound as possible, and the lack of incorporation of mechanistic insights, which in addition are often rather weak, must be of utmost concern to all toxicologists.

Toxicology is one of the few scientific disciplines that increase our knowledge in ways that might find major and often immediate public responses. The fears of chemical hazards to human health and the environment are important public issues—at least in our highly industrialized societies. The fear of chemicals is not new. As Pliny the Elder commented in the first century BC: "We use so many poisons to adjust the wine to our taste—and then we wonder why this does not do us any good." However, the massive "intrusion of chemistry into the human habitat" can no longer be denied. It forces us to confront the issue.

In future history books, the second half of the 20th century will be recorded as the era of synthetic chemistry. Beginning in the 1940s, we witnessed an explosive development of new industrial technologies and an integration of synthetic chemicals into our personal lives. In the United States alone, almost 600,000 chemical products are in use. About 70,000 of these products surround us daily, and the numbers are rising continuously. These substances have qualitatively and quantitatively different effects that may threaten our health. Initially, this type of chemistry was viewed positively as a benefit to our lives. Society has become more suspicious, however, due to wake-up calls such as Rachel Carson's book *The Silent Spring* (1962). Although people in the industrialized parts of the world enjoy much longer and qualitatively richer lives, the

public no longer sees science and technology only as valuable allies; rather, they view them, at least in part, as enemies of nature.

Public pressure has led to strict legislation with respect to chemical exposure, both in the environment and in the work place. Naturally, we all should welcome this development and demand that human welfare and the environment be protected. However, it is also uncontested that industry must be encouraged to invest in research to develop new materials and products that will be beneficial to all of us. The toxicological risk assessment—the evaluation of risks to human health and the environment from exposure to chemicals—is more and more a matter of sitting on the fence between these two postulates; it also stands on shaky scientific ground. On the one hand, chemical analytical techniques are capable of determining unimaginably small trace amounts of most chemicals, both synthetic and natural. The analytical results are often linked to symptoms induced by large doses of the respective compounds—without considering that analytical data and disease-related doses possibly differ by several orders of magnitude. Without such quantitative considerations of effect and risk assessment, however, the analytical findings are meaningless. On the other hand, toxicology does not always keep pace with advances in analytical technology. Despite huge gains in knowledge during the 1990s, it is unquestionable that, despite impressive progress, both our understanding of the mechanisms of toxicity and particularly our ability to assess risks quantitatively are very limited.

One can certainly quarrel over whether the increase in the use of chemicals—cosmetics, laundry detergents, food preservatives—is necessary, but this debate is irrelevant. These products are part of our lives and will stay with us into the forseeable future. The responsibility of toxicology is to ensure that humans are not subjected to unnecessary risks through exposure to these chemicals.

The foundation of toxicology is and will be for the forseeable future animal experiments that can point out potential hazards for the purpose of preventive health protection. For years, alternative methods have been advocated to replace intact animals by applying test systems other than intact organisms. However, these methods are valuable tools for special studies, but nothing more. The development of defined breeds of laboratory animals, as well as methods for their maintenance, has made it possible to carry out reproducible experiments. This must be viewed as the real beginning of experimental toxicology.

The most important principle of toxicology is the realization by Paracelsus that there are no toxic substances, only toxic dosages (applications) of compounds. In his third Carinthian Defence, Paracelsus made it quite clear that poison cannot be defined as a substance per se; rather, the same substance might be both poisonous and nonpoisonous: "Only the dose determines when a thing is not a poison." According to Paracelsus, we should define poison as a mechanism of action that is associated with both the chemical compound and the dose. Toxicologists are therefore confronted with the fact that—with exception of genotoxic compounds—any chemical is toxic above a certain threshold concentration (even 100 g of sodium chloride may be deadly), but not toxic at all below a threshold. This problem can only be solved by carefully analyzing the underlying mechanisms of action and the conditions that lead to toxicity. Threshold concentrations such as the no adverse effect level (NOAEL) are derived from toxicological investigations and, under the consideration of factors of (un)certainty, lead to human exposures that are tolerable from a toxicological point of view—that is, threshold values or guidelines such as the accepted daily intake (ADI).

The priority of modern toxicology is the detection of possible health hazards caused by the contamination of water, soil, air, and food. This poses especially difficult problems, as we are dealing here with biological activities of substances present in minute quantities. Initially, toxicology treated cases of acute poisoning after relatively high levels of exposure (Orfila, 1814). The clinical picture is characterized by more or less typical symptoms, a characteristic course of disease, and an obvious chronological relationship between intake of a compound and beginning of the disease; the effects are to a large extent reversible and the threshold values described above are applicable. Since the 1950s, toxicology has been mainly concerned with assessment and treatment of chronic intoxication and long-term effects of exogenous substances taken up in trace amounts. This development is due mainly to the identification of a new type of toxic effect, namely, the interaction of chemicals with genetic material.

The damage caused by these genotoxic effects—mutagenicity and carcinogenicity—is to a large extent irreversible and possibly occurs even with exposure to minute concentrations. The (often only theoretical) possibility that long-term exposure to minute doses might induce cancer has caused public fear and uncertainty. However, we all know very well that chemicals are acutely poisonous in high doses.

1. Five of 100,000 children die every year from accidental poisoning, especially from household chemicals. This risk is commonly accepted.
2. Most of our fellow citizens, however, are afraid of

a 1:1,000,000 risk (probability), if it exists at all, of developing cancer from exposure to asbestos in schools and demand a "cleanup," although it is exactly this measure that may cause the real asbestos threat.

Assessing the potential dangers of carcinogens is especially difficult. Presumably, there are multiple reasons that a normal cell may be transformed into a cancer cell. The underlying mechanisms of action are mostly unknown, and the test assays are inadequate in many ways. Nevertheless, there are certain basics of risk assessment. According to Paracelsus, exposure to concentrations below a threshold value is harmless. At present, such threshold values cannot be defined for carcinogens that interact with genetic material. Here, the experimentally unproven principle of stochastic effects is valid. It says that the probability of damage by a foreign compound declines with decreasing concentrations but will never be zero as long as a single molecule is still present. However, for a large number of nongenotoxic carcinogens in our environment—such as saccharin or some chlorinated hydrocarbons—there is no reason that we should not follow the rule of Paracelsus. We can no longer watch passively as it is declared reasonable that test results obtained with exorbitantly high dosages are extrapolated to the ppb or ppt range (1 part per billion or trillion) of normal environmental concentrations (comparable to a cube of sugar dissolved in Lake Geneva, Switzerland). If the one molecule hypothesis of carcinogenesis were correct, we all would be dead from cancer at an early age: Natural carcinogens occur normally in food at levels 10,000 times higher than those of residues of synthetic chemicals, and carcinogens are produced in food by ordinary cooking procedures. Furthermore, each cell normally contains elements that are carcinogenic, such as arsenic, cadmium, and uranium. Obviously, the toxicological assessment of both the chemicals and the specific situations in which humans are exposed to chemicals requires expert knowledge and great experience. The simple extrapolation of data from animal experiments to humans without considering the differences between species and mechanisms of action and without considering dose–response relationships inevitably leads to great uncertainties. "Extrapolation is more than using carbon paper" (A. F. Rahde). The assessment of the health risks from the exposure to chemicals must be based on established scientific knowledge. We can no longer tolerate irrationality when dealing with chemicals. Chemophobia, which is so much in the limelight, is often far from scientific reality: "Alarm is sounded long before it is clear what exactly we must be warned about—and those who paint the gloomiest pictures are trusted the most" (H. Rüdiger).

Humans are exposed to xenobiotics when using compounds with both beneficial and toxic effects (drugs) or to chemicals that have only toxic potentials. Xenobiotics generally show dose-dependent toxicity. The huge diversity in the use of and exposure to these potentially toxic xenobiotics has resulted in the cooperation of scientists from many disciplines with inevitably differing approaches. Thus, the study of toxicology can be divided into various fields, such as food, occupational medicine, environmental studies, and drug development.

Experimental toxicology collects data that are partly descriptive, but which also contribute to an understanding of the mechanism of the toxic effects observed. Thus, toxicological studies are performed not only in laboratories of the chemical and pharmaceutical industries, but also in research institutes concerned with biochemistry, physiology, pharmacology, cell biology, and pharmaceutical research.

Regulatory toxicology assesses the results of toxicological investigations with regard to the potential risk to humans and the environment. The existence of regulatory toxicology has resulted in the establishment of a large number of standardized or core study protocols. The quality of the data generated is maintained and guaranteed by the implementation of GLP (good laboratory practice) rules and the certification of the relevant laboratories by the responsible authorities. Standardization and the transparent documentation of the methods, animal strains, cell lines, and bacterial strains used are necessary to guarantee comparability of the data. Regulatory toxicology estimates the potential risk of exposure to xenobiotics of humans or the environment. Often only animal data are available for extrapolation to define an acceptable safety margin for public health. Risk represents the probability that a given xenobiotic will induce adverse effects or even severe damage; it is the reciprocal of safety. In other words, safety corresponds to the expectation of an acceptable risk under specified exposure conditions. The acceptable risk, however, must be defined on a case by case basis.

Thus, for example, an extremely toxic xenobiotic handled in a well-controlled manner will be less dangerous than a moderately toxic material released without appropriate control or assimilated in an uncontrolled manner. The assimilation of the material is a function of the exposure route and the physicochemical properties of the compound, which will have a major effect on the amount of compound absorbed and the exposure time. Toxicokinetic studies provide information about bioavailability, distribu-

tion to target organs, metabolism, and elimination. Investigation of the metabolism of a compound can give valuable insights into the mechanism of toxicity, such as in the case of the hepatotoxicity of acetaminophen or the porphyrinogenic effect of lead.

When toxicology is considered in depth, it becomes clear that the old concept of "toxicology is the pharmacology of high doses" is incomplete. In contrast, numerous actions and methods employed in toxicology have no relevance in pharmacology, examples being mutagenicity and carcinogenicity. These actions are, however, of the utmost importance in evaluating new chemical entities before their human exposure can be considered.

As a rule, the effects of xenobiotics are a function of dose, duration, and frequency of exposure. Toxicology discriminates among four categories of exposure, independent of the route of administration: (i) acute, (ii) subacute, (iii) subchronic, and (iv) chronic. *Acute* uptake indicates a single exposure lasting less than 24 hours, whereas the other categories describe repeated administrations. A *subacute* exposure describes repeated administrations of up to 28 days. *Subchronic* exposure reflects repeated administrations between 1 and 3 months. Exposure for more than 3 months is designated *chronic.*

Symptoms observed after repeated exposures may differ substantially from those observed after a single dose. For example, acute benzene exposure results in effects on the central nervous system, whereas leukemia is the prominent adverse event after prolonged exposure. Another well-known example is ethanol (ethyl alcohol, or alcohol), which after a single overdose mainly shows central nervous effects, for instance, disturbance of movements, euphoria, and unconsciousness, in addition to nausea, vomiting, and even death. Chronic abuse of alcohol results in liver cirrhosis, polyneurophathia, liver cancer, and addiction. Furthermore, symptoms resulting from acute toxic effects may be delayed, being observed only after several days, such as in the case of acetaminophen intoxication. This time lag can result in an acute intoxication being underestimated or even not being recognized by a practitioner.

Test strategies and requirements vary considerably according to the intended use of the compounds under consideration, whether for human use (e.g., drugs, food additives, and cosmetics) or bulk industrial production. Regulatory authorities have well-defined toxicological requirements for both classes of compounds.

Repeated or continuous exposure to xenobiotics may result in *tolerance,* that is, reduced effects due to prior exposure to the same or a chemically related compound. There are two possible causes of tolerance: the target organ is presented with a less toxic metabolite of the parent compound following modulation of metabolism, or the susceptibility of the target organ is reduced. The detailed mechanisms for these phenomena often need to be elucidated. Tolerance is observed for both drugs and other xenobiotics such as carbon tetrachloride. In the latter case, the trichloromethyl radical, the reactive metabolite responsible for hepatotoxicity, is formed to a lesser extent after repeated exposure. The heavy metal cadmium provides another example. Repeated uptake results in the induction of a cadmium-binding protein called metallothionein. As long as the concentration of free cadmium ions is practically zero, the symptoms of cadmium poisoning generally remain absent.

A further aspect of the toxicity of a compound is its *reversibility,* as shown by two classes of insecticides, carbamates and alkylated phosphates. For example, carbamate intoxication leads to reversible inhibition of acetylcholinesterase activity; however, the enzyme is reactivated after a few hours. In contrast, alkylated phosphates bind irreversibly to the same enzymes. The symptoms are identical in both intoxications. However, in the latter case, the acetylcholinesterase cannot be reactivated, and *de novo* erythropoesis is necessary. Consequently, the symptoms will persist much longer in the case of alkylated phosphate intoxication and therapeutic measures must be substantially different.

The *interaction* between two or more compounds can play an important role. Interaction can result in either an increase or a decrease in the toxic effects of a compound. A toxic effect is *additive* if, after simultaneous administration of two or more xenobiotics, it is the sum of the individual effects; this occurs, for example, after exposure to two different organophosphates. A *synergistic* effect will produce more severe symptoms than would be expected from the toxicities of the individual compounds; this is observed in the hepatotoxicity of carbon tetrachloride when it is given in the presence of ethanol. *Potentiation* occurs when a compound showing little or no toxicity markedly increases the toxicity of a second compound. As an example, isopropanol alone shows no hepatotoxicity; however, co-administration of isopropanol and carbon tetrachloride provokes hepatotoxic effects much more severe than would be expected from carbon tetrachloride alone.

Two compounds can interact to antagonize one another's effects. This phenomenon is the basis of *antidote* therapy. If two compounds counteract one another with regard to physiological effects, but not necessarily by the same mechanism of action, a *func-*

tional antagonism is involved. For example, several classes of compounds induce convulsions, which are antagonized by diazepam, but diazepam does not necessarily compete for the same binding site or receptor as the convulsant. *Chemical antagonism* occurs if a well-defined chemical reaction inactivates the toxic xenobiotic. Dimercaprol (British Anti-Lewisite, BAL) chelates arsenic, lead, and mercury ions, producing a nontoxic metal complex. By this simple chemical reaction, the chelating agent withdraws the toxic metal ions from the target organ and facilitates excretion. Competition for bonding to the same receptor is called *receptor antagonism;* an overdose of morphine can be antagonized by naloxone, and an overdose of benzodiazepam by flumenazil, as in each case both compounds are competing for the same receptor. In practice, however, the antagonist must be shown to have no deleterious pharmacodynamic effect. An antidote acting through a receptor antagonism does not necessarily compete with the xenobiotic directly. Atropine can be used to treat organophosphate poisoning, which causes accumulation of the neurotransmitter acetylcholine, by blocking the acetylcholine receptor.

As a result of developing strategies in drug research, such as combinatorial chemistry and high throughput screening (HTS), toxicology has developed an additional focus, involving mainly mechanistically oriented *in vitro* screening. These new pharmacological screening strategies allow the testing of several hundred thousand new chemical entities each year per laboratory. Consequently, an increasing demand for new test systems for use in early qualitative decision making in the development of new chemical entities can be expected. These screening systems will probably be focused mainly on mechanistically oriented toxicology with clearly defined target organs. *Predictive in vitro toxicology* should fulfill at least five criteria:

1. Target-related toxicity should correlate with the structure of the compounds and potential effects, for example, organ toxicity or metabolic effects.
2. Prediction of targets for epigenetic risk factors for carcinogenicity should be achieved.
3. A good correlation between results from *in vitro* and *in vivo* test systems should be demonstrated.
4. Test systems should have a low incidence of false negatives results.
5. An automatization of toxicological *in vitro* screening should be possible.

Much effort will be required to meet these goals. However, substantial progress is expected.

The responsibility that toxicology as a science has to humanity and the technological achievements of the chemical industry should encourage us to develop this discipline further, using great care and critical expertise. It is essential that any risk assessment be based on the best science available. The dramatic developments in toxicology and the substantial progress in cell and molecular biology must be integrated into toxicological risk assessment. Above all, the critical information must be communicated to the general public. More than ever before, physicians, chemists, and other scientists must become familiar with toxicological knowledge and the principles of toxicological safety assessment. This book is dedicated to this goal.

H. Marquardt, S. G. Schäfer, R. McClellan, F. Welsch

This Page Intentionally Left Blank

Introduction

BRUCE N. AMES, LOIS SWIRSKY GOLD

Division of Biochemistry and Molecular Biology and National Institutes of Environmental Health Sciences Center, University of California at Berkeley, Berkeley, California

In this introductory essay we raise four issues in the fuzzy interface between toxicology, nutrition, public health, and government regulations.

PARACELSUS TO PARASCIENCE: THE DOSE/TRACE MAKES THE POISON

About 50% of all chemicals—whether natural or synthetic—that have been tested in standard, high-dose animal cancer tests are rodent carcinogens (Table 1). What are the explanations for this high percentage? In standard cancer tests, rodents are given a chronic, near-toxic dose: the maximum tolerated dose (MTD). Evidence is accumulating that cell division caused by the high dose itself, rather than the chemical per se, can contribute to cancer in these tests. High doses can cause chronic wounding of tissues, cell death, and the consequent chronic cell division of neighboring cells, which is a risk factor for cancer. Each time a cell divides, the probability increases that a mutation will occur, thereby increasing the risk for cancer. At the low levels to which humans are usually exposed, such increased cell division does not occur. In addition, tissues injured by high doses of chemicals have an inflammatory immune response involving the activation of white cells in response to cell death. Activated white cells release mutagenic oxidants (including peroxynitrite, hypochlorite, and hydrogen peroxide). Therefore, the very low levels of chemicals to which humans are exposed through water pollution or synthetic pesticide residues may pose no or minimal cancer risks.

Is the high positivity rate due to selecting more suspicious chemicals to test, which is a likely bias because cancer testing is both expensive and time-consuming, and it is prudent to test suspicious compounds? One argument against selection bias is the high positivity rate for drugs (Table 1) because drug development tends to select chemicals that are not mutagens or expected carcinogens. A second argument against selection bias is that the knowledge needed to predict carcinogenicity in rodent tests is highly imperfect, even now after decades of testing results have become available on which to base predictions. For example, a prospective prediction exercise was conducted by several experts in 1990 in advance of the 2-year NTP bioassays. There was a wide disagreement among them on which chemicals would be carcinogenic when tested and the level of accuracy varied by expert, thus indicating that predictive knowledge is highly uncertain.

It seems likely that a high proportion of all chemicals, whether synthetic or natural, might be "carcinogens" if administered in the standard rodent bioassay at the maximum tolerated dose, primarily due to the effects of high doses on cell division and DNA damage. Without additional data on how a chemical causes cancer, the interpretation of a positive result in a rodent bioassay is highly uncertain. The induction of cancer could be the result of the high doses tested.

In regulatory policy, the virtually safe dose (VSD), corresponding to a maximum hypothetical risk of one cancer in a million, is estimated from bioassay results using a linear model that assumes cancer causation is directly proportional to dose and that there are no unique effects of high doses. To the extent that carci-

TABLE 1 Proportion of Chemicals Evaluated as Carcinogenic

	Proportion	Percent
Chemicals tested in both rats and mice[a]	330/559	(59%)
Naturally occurring chemicals	73/127	(57%)
Synthetic chemicals	257/432	(59%)
Chemicals tested in rats and/or mice[a]		
Chemicals in Carcinogenic Potency Database	668/1275	(52%)
Natural pesticides	35/64	(55%)
Mold toxins	14/23	(61%)
Chemicals in roasted coffee	19/28	(68%)
Innes negative chemicals retested[a,b]	16/34	(47%)
Physician's Desk Reference (PDR)		
Drugs with reported cancer tests[c]	117/241	(49%)
FDA database of drug submissions[d]	125/282	(44%)

[a]From Gold, L. S. *et al.* (1997).

[b]The 1969 study by Innes *et al.* is frequently cited as evidence that the proportion of carcinogens is low, as only 9% of 119 chemicals tested (primarily pesticides) were positive. However, these tests, which were only in mice with few animals per group, lacked the power of modern tests. Of the 34 Innes negative chemicals that have been retested using modern protocols, 16 were positive.

[c]Davies and Monro (1995).

[d]Contrera *et al.* (1997). 140 drugs are in both the FDA and PDR databases.

nogenicity in rodent bioassays is due to the effects of high doses for the nonmutagens and a synergistic effect of cell division at high doses with DNA damage for the mutagens, then this model is inappropriate.

EVEN RACHEL CARSON WAS MADE OF CHEMICALS: NATURAL VERSUS SYNTHETIC CHEMICALS

About 99.9% of the chemicals humans ingest are natural. The amounts of synthetic pesticide residues in plant foods are insignificant compared to the amount of natural pesticides produced by plants themselves. Of all dietary pesticides that humans eat, 99.99% are natural: they are chemicals produced by plants to defend themselves against fungi, insects, and other animal predators. Each plant produces a different array of such chemicals.

On average, Americans ingest roughly 5000 to 10,000 different natural pesticides and their breakdown products. Americans eat about 1500 mg of natural pesticides per person per day, which is about 10,000 times more than they consume of synthetic pesticide residues.

Even though only a small proportion of natural pesticides have been tested for carcinogenicity, 35 of the 63 tested are rodent carcinogens. Naturally occurring pesticides that are rodent carcinogens are ubiquitous in fruits, vegetables, herbs, and spices (Table 2).

Cooking foods produces about 2000 mg per person per day of burnt material that contains many rodent carcinogens and many mutagens. By contrast, the residues of 200 synthetic chemicals measured by FDA, including the synthetic pesticides thought to be of greatest importance, average only about 0.09 mg per person per day. In a single cup of coffee, the natural chemicals that are known rodent carcinogens are about equal in weight to a year's worth of synthetic pesticide residues that are rodent carcinogens, even though only 3% of the natural chemicals in roasted coffee have been adequately tested for carcinogenicity (Table 3). This does not mean that coffee or natural pesticides are dangerous but rather that assumptions about high-dose animal cancer tests for assessing human risk at low doses need reexamination. No diet can be free of natural chemicals that are rodent carcinogens.

Gaining a broad perspective about the vast number of chemicals to which humans are exposed can be helpful when setting research and regulatory priorities. Rodent cancer tests provide little information about how a chemical causes cancer or about its effects at a low dose. The assumption that synthetic chemicals are hazardous has led to a bias in testing, such that synthetic chemicals account for 77% (432 of 559) of the chemicals tested chronically in both rats and mice (Table 1). The natural world of chemicals has never been tested systematically. One reasonable strategy is to use a rough index to compare and rank possible carcinogenic hazards from a wide variety of

TABLE 2 Carcinogenicity in Rodents of Natural Plant Pesticides Tested[a]

Carcinogens (N = 35)

Acetaldehyde methylformylhydrazone, Allyl isothiocyanate, Arecoline.HCl, Benzaldehyde, Benzyl acetate, Caffeic acid, Catechol, Clivorine, Coumarin, Crotonaldehyde, Cycasin and methylazoxymethanol acetate, 3,4-dihydrocoumarin, Estragole, Ethyl acrylate, N^2-γ-Glutamyl-*p*-hydrazinobenzoic acid, Hexanal methylformylhydrazine, *p*-Hydrazinobenzoic acid.HCl, Hydroquinone, 1-Hydroxyanthraquinone, Lasiocarpine, D-Limonene, 8-Methoxypsoralen, *N*-Methyl-*N*-formylhydrazine, α-Methylbenzyl alcohol, 3-Methylbutanal methylformylhydrazone, Methylhydrazine, Monocrotaline, Pentanal methylformylhydrazone, Petasitenine, Quercetin, Reserpine, Safrole, Senkirkine, Sesamol, Symphytine

Noncarcinogens (N = 28)

Atropine, Benzyl alcohol, Biphenyl, D-Carvone, Deserpidine, Disodium glycyrrhizinate, Emetine.2HCl, Ephedrine sulphate, Eucalyptol, Eugenol, Gallic acid, Geranyl acetate, β-*N*-[γ-l(+)-Glutamyl]-4-hydroxy-methylphenylhydrazine, Glycyrrhetinic acid, *p*-Hydrazinobenzoic acid, Isosafrole, Kaempferol, DL-Menthol, Nicotine, Norharman, Pilocarpine, Piperidine, Protocatechuic acid, Rotenone, Rutin sulfate, Sodium benzoate, Turmeric oleoresin, Vinblastine

Rodent carcinogen sources

Absinthe, Allspice, Anise, Apple, Apricot, Banana, Basil, Beet, Broccoli, Brussels sprouts, Cabbage, Cantaloupe, Caraway, Cardamom, Carrot, Cauliflower, Celery, Cherries, Chili pepper, Cinnamon, Cloves, Cocoa, Coffee, Collard greens, Comfrey herb tea, Coriander, Corn, Currants, Dill, Eggplant, Endive, Fennel, Garlic, Grapefruit, Grapes, Guava, Honey, Honeydew melon, Horseradish, Kale, Lemon, Lentils, Lettuce, Licorice, Lime, Mace, Mango, Marjoram, Mint, Mushrooms, Mustard, Nutmeg, Onion, Orange, Paprika, Parsley, Parsnip, Peach, Pear, Peas, Black pepper, Pineapple, Plum, Potato, Radish, Raspberries, Rhubarb, Rosemary, Rutabaga, Sage, Savory, Sesame seeds, Soybean, Star anise, Tarragon, Tea, Thyme, Tomato, Turmeric, Turnip

[a]Fungal toxins are not included. Reprinted from Gold *et al.*, In *Food Chemical Risk Analysis* (D. Tennant, ed.), pp. 267–295, © 1997 Aspen Publishers, Inc.

chemical exposures at levels that humans typically receive, and then to focus on those that rank highest. Ranking is a critical first step that can help to set priorities for selecting chemicals for long-term cancer tests, studies on mechanism, epidemiological research, and regulatory policy. Although one cannot say whether the ranked chemical exposures are likely to be of major or minor importance in human cancer, it is not prudent to focus attention on the possible hazards at the bottom of a ranking if, using the same methodology to identify hazard, there are numerous common human exposures with much greater possible hazards. Our analyses are based on the HERP index (Human Exposure/Rodent Potency), which indicates what percentage of the rodent carcinogenic potency a human receives from a given daily lifetime exposure. A ranking based on standard regulatory risk assessment would be similar.

TABLE 3 Carcinogenicity in Rodents of Natural Chemicals in Roasted Coffee[a]

Positive (N = 19)

Acetaldehyde, Benzaldehyde, Benzene, Benzofuran, Benzo[*a*]pyrene, Caffeic acid, Catechol, 1,2,5,6-Dibenzanthracene, Ethanol, Ethylbenzene, Formaldehyde, Furan, Furfural, Hydrogen peroxide, Hydroquinone, Limonene, Styrene, Toluene, Xylene

Not positive (N = 8)

Acrolein, Biphenyl, Choline, Eugenol, Nicotinamide, Nicotinic acid, Phenol, Piperidine

Uncertain: Caffeine

Yet to test: ~1000 chemicals

[a]Gold *et al.* (1997).

Overall, the analyses have shown that HERP values for some historically high exposures in the work place (e.g., butadiene and tetrachloroethylene) and some pharmaceuticals (e.g., clofibrate) rank high, and that there is an enormous background of naturally occurring rodent carcinogens in the average consumption of common foods that casts doubt on the relative importance of low-dose exposures to residues of synthetic chemicals such as pesticides. A committee of the National Research Council of the National Academy of Sciences, recently reached similar conclusions about natural vs synthetic chemicals in the diet and called for further research on natural chemicals.

The possible carcinogenic hazards from synthetic pesticides are minimal compared to the background of nature's pesticides, though neither may be a hazard at the doses consumed. Analysis also indicates that many ordinary foods would not pass the regulatory criteria used for synthetic chemicals. Caution is necessary in drawing conclusions from the occurrence in the diet of natural chemicals that are rodent carcinogens. It is not argued here that these dietary exposures are necessarily of much relevance to human cancer. The data call for a reevaluation of the utility of animal cancer tests in protecting the public against minor hypothetical risks.

It is often assumed that because natural chemicals are part of human evolutionary history, whereas synthetic chemicals are recent, the mechanisms that have evolved in animals to cope with the toxicity of natural chemicals will fail to protect against synthetic chemicals. This assumption is flawed for several reasons:

1. Humans have many natural defenses that buffer against normal exposures to toxins and these are usually general, rather than tailored for each specific chemical. Thus they work against both natural and synthetic chemicals. Examples of general defenses include the continuous shedding of cells (the surface layers of the mouth, esophagus, stomach, intestine, colon, skin, and lungs are discarded every few days); DNA-repair enzymes, which repair DNA that was damaged from many different sources; and detoxification enzymes of the liver and other organs, which generally target classes of chemicals rather than individual chemicals. That human defenses are usually general, rather than specific for each chemical, makes good evolutionary sense. The reason that predators of plants evolved general defenses is presumably to be prepared to counter a diverse and ever-changing array of chemicals in plants in an evolving world; if a herbivore had defenses against only a set of specific plant pesticides, it would be at a great disadvantage in obtaining new food when favored foods became scarce or evolved new chemical defenses.

2. Various natural toxins, which have been present throughout vertebrate evolutionary history, nevertheless cause cancer in vertebrates. Mold toxins, such as aflatoxin, have been shown to cause cancer in rodents and other species, including humans (Table 1). Many of the common elements are carcinogenic to humans at high doses (e.g., salts of cadmium, beryllium, nickel, chromium, and arsenic) despite their presence throughout evolution. Furthermore, epidemiological studies from various parts of the world show that certain natural chemicals in food may be carcinogenic risks to humans; for example, the chewing of betel nuts with tobacco caused oral cancer.

3. Humans have not had time to evolve a "toxic harmony" with all of their dietary plants. The human diet has changed markedly in the last few thousand years. Indeed, very few of the plants that humans eat today (e.g., coffee, cocoa, tea, potatoes, tomatoes, corn, avocados, mangoes, olives, and kiwi fruit) would have been present in a hunter–gatherer's diet. Natural selection works far too slowly for humans to have evolved specific resistance to the chemicals in these newly introduced plants.

4. DDT is often viewed as the typically dangerous synthetic pesticide because it concentrates in tissues and persists for years, being slowly released into the bloodstream. DDT, the first synthetic pesticide, eradicated malaria from many parts of the world, including the United States. It was effective against many vectors of disease such as mosquitoes, tsetse flies, lice, ticks, and fleas. DDT was also lethal to many crop pests and significantly increased the supply and lowered the cost of food, making fresh nutritious foods more accessible to poor people. DDT was also of low toxicity to humans. A 1970 National Academy of Sciences report concluded: "In little more than two decades DDT has prevented 500 million deaths due to malaria, that would otherwise have been inevitable." There is no convincing epidemiological evidence, nor is there much toxicological plausibility, that the levels of DDT normally found in the environment are likely to be a significant contributor to human cancer. DDT was unusual with respect to bioconcentration and, because of its chlorine substituents, it takes longer to degrade in nature than most chemicals; however, these are properties of relatively few synthetic chemicals. In addition, many thousands of chlorinated chemicals are produced in nature and natural pesticides also can bioconcentrate if they are fat soluble. Potatoes, for example, naturally contain the fat-soluble neurotoxins solanine and chaconine, which can be detected in the bloodstream of all potato eaters. High levels of these potato neurotoxins have been shown to cause birth defects in rodents.

5. Since no plot of land is immune to attack by insects, plants need chemical defenses—either natural or synthetic—to survive pest attack. Thus, there is a trade-off between naturally occurring pesticides and synthetic pesticides. One consequence of the disproportionate concern about synthetic pesticide residues is that some plant breeders develop plants to be more insect-resistant by making them higher in natural toxins. A recent case illustrates the potential hazards of this approach to pest control. When a major grower introduced a new variety of highly insect-resistant celery into commerce, people who handled the celery developed rashes when they were subsequently exposed to sunlight. Some detective work found that the pest-resistant celery contained 6200 parts per billion (ppb) of carcinogenic (and mutagenic) psoralens instead of the 800 ppb present in common celery.

ERRORS OF OMISSION: TOXICITY OF INSUFFICIENT MICRONUTRIENTS

Hormonal factors, dietary imbalances, infection and inflammation, and genetic factors, none of which involve a carcinogenic chemical, are major contributors to cancer.

TABLE 4 Review of Epidemiological Studies on Cancer Showing Protection by Consumption of Fruits and Vegetables[a]

Cancer site	Fraction of studies showing significant cancer protection	Median relative risk of low-quarter vs. high-quarter consumption
Epithelial		
Lung	24/25	2.2
Oral	9/9	2.0
Larynx	4/4	2.3
Esophagus	15/16	2.0
Stomach	17/19	2.5
Pancreas	9/11	2.8
Cervix	7/8	2.0
Bladder	3/5	2.1
Colorectal	20/35	1.9
Miscellaneous	6/8	—
Hormone-dependent		
Breast	8/14	1.3
Ovary, endometrim	3/4	1.8
Prostate	4/14	1.3
Total	129/172	

[a]Block *et al.* (1992).

The high consumption of fruits and vegetables is associated with a lower risk of degenerative diseases including cancer, cardiovascular disease, cataracts, and brain dysfunction. Over 200 epidemiological studies have been reviewed that show, with great consistency, an association between the low consumption of fruits and vegetables and cancer incidence (Table 4).

The quarter of the population with the lowest dietary intake of fruits and vegetables has roughly twice the cancer rate of the quarter with the highest intake for most types of cancer (lung, larynx, oral cavity, esophagus, stomach, colorectal, bladder, pancreas, cervix, and ovary). In the United States, 80% of children and adolescents and 68% of adults did not meet the intake recommended by the National Cancer Institute and the National Research Council of five servings of fruits and vegetables per day.

Publicity about hundreds of minor hypothetical risks, such as pesticide residues, can result in a loss of perspective about what is important: half the public does not know that fruit and vegetable consumption is a major protection against cancer. Fruits and vegetables are of major importance for reducing cancer; if they become more expensive because of reduced use of synthetic pesticides, then consumption will decline and cancer is likely to increase. People with low incomes eat fewer fruits and vegetables and spend a higher percentage of their income on food.

Folic acid deficiency, one of the most common vitamin deficiencies among people consuming few dietary fruits and vegetables, causes extensive chromosome breaks in human genes. Approximately 10% of the U.S. population has a folate level lower than that at which chromosome breaks occur. In two small studies of low-income (mainly African-American) elderly and adolescents done nearly 20 years ago, nearly half had folate levels this low although the issue should be reexamined. The rate of chromosome breaks in humans is reduced by folate administration. Chromosome breaks could contribute to the increased risk of cancer and cognitive defects associated with folate deficiency in humans. Folate deficiency also damages human sperm, causes neural tube defects in the fetus, and causes about 10% of U.S. heart disease. Diets deficient in fruits and vegetables are commonly low in folate, antioxidants (e.g., vitamin C), and many other micronutrients, and result in DNA damage and higher cancer rates.

Antioxidants, such as vitamin C (whose dietary source is fruits and vegetables), vitamin E, and selenium, protect against oxidative damage from normal metabolism, smoking, and inflammation. Because radiation causes oxidative damage, insufficiency of dietary antioxidants is a radiation mimic.

Deficiency of any one of nine dietary micronutrients—folic acid, niacin, iron, zinc, selenium, and vitamins B6, B12, C, and E—appear to act as radiation

mimics in breaking DNA and chromosomes or causing oxidative damage to DNA or both. Some of these micronutrients come from fruits and vegetables and could account for a good part of their protective effect against cancer.

Many other micronutrients whose main dietary sources are other than fruits and vegetables, also are likely to play a significant role in the prevention and repair of DNA damage, and thus are important to the maintenance of long-term health. Deficiency of vitamin B12 (found in about 10% of U.S. elderly) causes a functional folate deficiency, the accumulation of homocysteine, a risk factor for heart disease, and probably chromosome breaks. Strict vegetarians are at increased risk of developing a vitamin B12 deficiency. Niacin contributes to the repair of DNA breaks. As a result, dietary insufficiencies of niacin (15% of some populations are deficient), folate, and antioxidants may act together to increase DNA damage. There is also evidence that low intake (<50% recommended daily allowance) of zinc, iron, selenium, or vitamin B6 (7–20% of the U.S. population is low in each) can lead to DNA damage. Half of the U.S. population may be low in at least one of these nine micronutrients. Optimizing micronutrient intake (through better diets, fortification of foods, or multivitamin–mineral pills) can have a major impact on health at a low cost. Increasing research in this area and efforts to increase micronutrient intake and more balanced diets should be high priorities for public policy.

DAMAGE BY DISTRACTION: REGULATING LOW HYPOTHETICAL RISKS

Synthetic hormone mimics have become an environmental issue. Hormonal factors are important in cancer, as mentioned above. The book *Our Stolen Future* (Colburn *et al.*, 1996) claims that traces of synthetic chemicals, such as pesticides with weak hormonal activity, may contribute to cancer and reduce sperm counts. The book ignores the fact that our normal diet contains natural chemicals that have estrogenic activity millions of times higher than that due to the traces of synthetic estrogenic chemicals and that lifestyle factors can markedly change the levels of endogenous hormones. The low levels of human exposure to residues of industrial chemicals are toxicologically implausible as a significant cause of cancer or of reproductive abnormalities, especially when compared to the natural background. In addition, it has not been shown convincingly that sperm counts are declining, and even if they were there are many more likely causes, such as smoking and diet.

Because there is no risk-free world and resources are limited, society must set priorities based on benefits and cost-effectiveness in order to save the most lives. The EPA projected in 1991 that the cost to society of U.S. environmental regulations in 1997 would be about $140 billion per year (about 2.6% of the gross national product). Most of this cost is to the private sector. Several economic analyses by others have concluded that current expenditures are not cost-effective; that is, resources are not being used so as to save the most lives per dollar. One estimate is that the United States could prevent 60,000 deaths per year by redirecting the same dollar resources to more cost-effective programs. For example, the median toxin-control program costs 146 times more per year of life saved than the median medical-intervention program. This difference is likely to be greater because cancer-risk estimates for toxin-control programs are worst-case hypothetical estimates, and the true risks at low dose are often likely to be zero. Rules regulating air and water pollution are necessary (e.g., it was a public health advance to phase lead out of gasoline) and clearly cancer prevention is not the only reason for regulations. However, worst-case assumptions in risk assessment represent a policy decision, not a scientific one, and they confuse attempts to allocate money effectively for public health.

Regulatory efforts to reduce low-level human exposures to synthetic chemicals because they are rodent carcinogens are expensive; they aim to eliminate minuscule concentrations that now can be measured with improved techniques. These efforts are distractions from the major task of improving public health through increasing scientific understanding of how to prevent cancer (e.g., what aspects of diet are important), increasing public understanding of how lifestyle influences health, and improving our ability to help individuals alter their lifestyles.

Why has the government focused on minor hypothetical risks at huge cost? A recent article in *The Economist* (December 20, 1997) had fairly harsh judgment:

> Predictions of ecological doom, including recent ones, have such a terrible track record that people should take them with pinches of salt instead of lapping them up with relish. For reasons of their own, pressure groups, journalists and fameseekers will no doubt continue to peddle ecological catastrophes at an undiminishing speed . . . . Environmentalists are quick to accuse their opponents in business of having vested interests. But their own incomes, their fame and their very existence can depend on supporting the most alarming versions of every environmental scare. 'The whole aim of practical politics,' said H.L. Mencken, 'is to keep the popu-

lace alarmed—and hence clamorous to be led to safety—by menacing it with a series of hobgoblins, all of them imaginary.' Mencken's forecast, at least, appears to have been correct.

Aaron Wildavsky discusses worst-case risk assessment in his book *But Is It True: A Citizen's Guide to Environmental Health and Safety Issues.*

> We should be guided by the probability and extent of harm, not by its mere possibility. The search for possibilities is endless and it trivializes the subject. There is bound to be great diversion of resources without reducing substantial sources of harm. Consternation is created but health is not enhanced . . . . Weak causes are likely to have weak effects. Our search should be for strong causes with palpable effects, like cigarette smoking. They are easier to find and their effects are much more important to control . . . . The past necessity of proving harm has been replaced by a reversal of causality: now the individuals and businesses must prove that they will do no harm. My objection to this . . . is profound: our liberties are curbed and our health is harmed.

Acknowledgments

This essay has been adapted in part from Ames and Gold, *FASEB J.* **11,** 1041–1052 (1997) and Ames, *Toxicol. Lett.* **102/103,** 5–18 (1998) with permission from Elsevier Science. For detailed literature, the reader is referred to both publications.

Bibliography

Block, G. *et al.* (1992). *Nutr. and Cancer* **18,** 1–29.

Contrera, J., *et al.* (1997). *Regul Toxicol. Pharmacol.* **25,** 130–145.

Davies, T., and Monro, A. (1995). *J. Am. Coll. Toxicol.* **14,** 90–107.

Gold, L. S., *et al.* (1997). In *Handbook of Carcinogenic Potency and Genotoxicity Databases* (L. S. Gold and E. Zeiger, eds.) pp. 661–685. CRC Press, Boca Raton, Florida.

Gold, L. S., *et al.* (1997a). In *Food Chemical Risk Analysis* (D. Tennant, ed.), pp. 267–295. Aspen Publishers, Gaithersburg, Maryland.

Innes, J., *et al.* (1969). *J. Natl. Cancer Inst.* **42,** 1101–1114.

This Page Intentionally Left Blank

PART

I

FUNDAMENTALS

This Page Intentionally Left Blank

CHAPTER

1

History of Toxicology

DIETRICH MILLES
Center for Social Policy Research, University of Bremen, Bremen, Germany

It has been known for centuries that substances affect the human body. From earliest times, people have wanted to know which substances support which bodily functions, whether they calm or paralyze them, lead to healing or death, and how they do so. It was established that the effects of a substance depended on its state: liquid or powder, in work or building materials or in foodstuffs, or through radiation. A long-lasting argument flared up as to whether substances are inherently effective or only in specific quantities. There was research into why different substances work differently on different people, in different environments, in different combinations. History supplies us not only with much evidence about the most important problems of toxicology, which vary with time, but also about the continuing efforts to approach its tasks in a responsible way, both scientifically and practically. The following introduction can deal with only some of the more important endeavors.

TOXICOLOGY—THE SCIENCE OF POISON

The word "toxicology" comes from the Greek word for poison (*toxicon*) and scientific study (*logos*) and was coined in the 17th century. To begin with, toxicology developed as an empirical science linked to medicine. The great turning point for the science came through Paracelsus, who worked on chemical processes in life functions and who emphasized the importance of the quantity of a substance: "Was ist das Allein die Dosis macht, daß ein ding gifft ist" ("Is there anything that is not poison? Everything is poison, and nothing is without poison. The dose alone makes a thing poisonous").

Toxicology was originally an empirical science, and did not evolve into a volumetric science until the emergence of chemistry and analytical science. The disciplines of pharmacy, chemistry, and toxicology branched off from medicine with the development of methods. At the University of Leiden in the Netherlands, Hermann Boerhaave, in a pathbreaking move, combined the chairs of medicine, botany, and chemistry from 1718 to 1729, while H. M. Bracken at the University of Minnesota subsumed under the subject of medical remedies ("materia medica"), phramacognosy, pharmacology, toxicology, and pharmacy in 1859. Towards the end of the 19th century, the disciplines of pharmacology and toxicology, linked until then, were separated due to the growth of the chemical industry and chemical products.

The distinct object of scientific inquiry in the ancient and the medieval world was "pharmakon," and there was no differentiation between cure and poison. In more modern times, we find in Zedler's universal encyclopedia (1735), under the heading "poison," "that which is very damaging, if not necessarily fatal, to the human body, when body parts come into contact with it, internally or externally." Louis Lewin, a distinguished chronicler in this field, differentiated in 1907 among four main types of poison: (1) inflammatory or caustic poisons, (2) poisons causing metabolic disturbance, (3) poisons of the nervous system, and (4) poisons of the blood. Through this differentiation

of the types of poison, he attempted to define the subject matter of toxicology as regards causes and effects.

Leading figures who from the end of the 19th century onwards defined the outlines in the field of toxicology typified the close relationships that extended beyond the boundaries of the subject. Lewin belonged to this generation (as did K. B. Lehmann, J. Abel, R. Kobert, T. H. Legge, L. Teleky, and others), who, at an intersection between medicine, pharmacology, pharmacy, chemistry, ergonomics, and technology, combined exact research into effect through analysis of cause and preventative reasoning. Thus, the Zurich toxicologist and forensic medical expert Heinrich Zangger, with the chlorine-gas attacks of the First World War still fresh in his mind, demanded in 1920 that guaranteed levels of safety be achieved through scientific identification of risk and through statistical examination: "Risk and its scientific control must remain in step." Medicine could then encourage and keep pace with technological advancement without undue risk, in particular without "scientifically establishable, chronic, permanent, unseen, degenerative damage to mankind." Meeting this challenge, toxicology has taken up its place in the ranks of the acknowledged sciences.

EARLY FUNDAMENTAL DEVELOPMENTS

The use of poisons and active substances dates from the earliest days of human evolution. Primitive man certainly instinctively avoided poisonous substances just as animals do and particularly feared the bite of poisonous animals. From such an animal instinct we can also infer the use of remedies taken internally; naturally occurring substances could be taken to help maintain physical and spiritual balance.

Primitive peoples used laxatives, emetics, enemas, fumigants, medicinal baths, and sudorifics. Narcotics are among the first plant-based active substances for which there is evidence. The first collections of information about the effects of compounds from various plants (among them aconite, arsenic, and opium) were made in Persia and China. The Ebers Papyrus lists over 800 substances used in Egypt 2000 years ago. In the 16th century B.C. opium, hyoscyamus (henbane), scilla, aloe, mint, calamus, juniper, cassia, sandalwood, myrrh, styrax, wine, honey, antimony, verdigris, iron, and animal organs, and secretions were used as medical remedies. Reports from the Trojan War (12th century B.C.) mention the use of wine laced with opium for the wounded.

The conditions for recovery or harm attributed to remedies and poisons lay beyond people's lives. The particular effects of substances could be easily understood as mystical concepts. "Harmful magic" was used to explain symbolic poisons and remedies, and "foreign bodies" their effects. This kind of understanding is illustrated by their attitudes toward plants surrounded in mystery, like mandrake (containing the alkaloids scopolamine, atropine, and hyoscyamine), which was as famous as it was poisonous; the same applies to opium poppies and hemlock.

Within the framework of primitive magic ideas of taboo, the concepts of poison and sin fused together. Both of these led to contagion and poisoning. Through efforts to purge sin, through cleansing such as blood-letting, the attempt to combat disease finally became a rational and logical act. In ancient Greece there was no strict dividing line between poisons and cures. Instead, an additional classification, such as "harmful" or "fatal," was necessary. Not until the writings of Hippocrates is there a stricter differentiation between *pharmakon* (collections of herbal or plant-based remedies, e.g., anaesthetizing sponges) and foodstuffs.

Significant experiments with medical remedies were made by Heraclides of Tarentum (1st century B.C.). At this time various minor tyrants were extremely interested in research into poisons and their antidotes, which could put the life of a dictator at risk or save it. The most famous of these was King Mithridates VI of Pontus (132–63 B.C.), who not only stood up to the Romans, but who tested out poisons on condemned criminals and tried to achieve immunity for himself by repeated small doses. An antidote named after him (Mithridatium, an electuary) was still being used in the Middle Ages. Important discoveries were made by Pedanios Dioscorides (64 A.D.) who knew lead chloride (Chlorblei), zinc oxide, sodium carbonate, sulphide of arsenic, and indigo, and who described the process by which mercury is obtained from cinnabar and the manufacture of lead acetate, limewater, and copper sulphate.

But there was another specialized use for toxic properties: In Athens, drinking the juice of hemlock was chosen as the death penalty for political criminals. In the 16th century we find aconite being used for an execution in the presence of Pope Clement VII. The most horrific effects of poison (Zyklon B) were seen in the gas chambers of the German concentration camps. Even today in some states of the U.S. (in Utah and Delaware, for example), the death penalty is carried out with lethal injections, in others with gas.

Greek doctors, whom the Romans at first denounced as quacks, disseminated medical knowledge

throughout the Roman Empire with the support of Julius Caesar. The earliest preserved compilation on the subject of medical remedies from the first century A.D. handed down to us over 1000 remedies. Particularly worthy of mention is Galenos of Pergamon (129–199 A.D.), who in his comprehensive review of medical knowledge in the ancient world also described pharmaceutical preparations in the form of tinctures, extracts, ointments, and so on, which were kept in stock. His doctrines were based on the theory of the pathology of the humors, according to which illnesses are caused by evil fluids, the humors, and can be treated according to the principle of opposites as well as by cleansing of the body, mainly through enemas and blood-letting. His theories were fundamental, too, to monastic medicine and remained unchanged until the advent of Paracelsus.

FROM ALCHEMY TO THE DEVELOPMENT OF A PURE SCIENCE

The early history of toxicology can be written as the history of pharmaceutics, but also as the history of forensic medicine. Priests and rulers, villains and assassins—all used poisons and were interested in the effects of poison from the point of view of perpetrator and victim. Sure knowledge and the correct diagnosis were, however, dependent on detection of the poison. But the examination of corpses was bound up in ancient times with religious and mystic taboos; the *corpus iuris civilis* (439 A.D.) did not permit obduction. It was not until the 15th century that the Catholic church allowed postmortem examinations for scientific purposes; in practice the taboos mentioned are still in effect today. As the numbers of murders by poisoning and the pursuit of the, mainly female, poisoners increased, and citizens increasingly felt the need to feel secure, there finally grew too the public desire for cases of poisoning to be cleared up.

To begin with, substances were used in many combinations based on Greek and Arab traditions—this continued into the Middle Ages. Along with henbane, mandrake, opium, and hemlock, alcohol gained in importance. Early mixtures consisted of powdered plants and heavy metals, in particular mercury and arsenic. The recipes for these mixtures were handed down to more recent times and took on new significance in the battle against syphilis, which spread through Europe like an epidemic after the return of Christopher Columbus. A further contributor to its spread was the army of the French king, Charles VIII. The treatment of the condition was as disastrous as it was ineffective. Arsenic compounds played a leading pharmaceutical role for a long time. Even Hippocrates recommended a paste containing arsenic sulfide in the treatment of excrescences. The infamous Aqua Tofana (white arsenic boiled with lead and antimony), was in its time a favorite with murderers.

Alchemy, which held sway in the Middle Ages and which came originally from Arabia, was strongly linked to metallurgy. It was thought that mercury could be changed into gold using the "philosopher's stone," and "elixirs of life" were concocted from heavy metals. In the course of such experimentation tin chloride was discovered in 1604 by Andreas Libavius. In 1662 tartar emetic was discovered by Adrian v. Mynsicht, and at the beginning of the same century Johann R. Glauber (1604–1670) discovered Glauber's salt. He named the sodium sulfate, which he was the first to create, "sal mirabile"—the wonder salt.

The transition to the chemical synthesis of effective medical remedies took place in the Renaissance and was linked to Theophrastus Bombastus von Hohenheim (1493–1541), known as Paracelsus. His work was not based on juices and their balance, but on effective natural elements and analogous chemical principles: salt (solid, permanent), sulfur (inflammable), quicksilver (pertaining to Mercury—changeable, active). Paracelsus recommended treating the sick body with compensatory metal salts such as sublimate, lead acetate, silver nitrate, arsenic sulfide, and also salts of antimony, bismuth, zinc, and gold. Antimony in particular was vehemently argued over in the 16th and 17th centuries. In disputes of this kind, people eventually became convinced that it was feasible to use artificially manufactured substances for medical purposes. Paracelsus viewed the human organism as a kind of laboratory that took in food and other substances, processed them, and excreted them. Extracts and tinctures were therefore investigated to see how they would work in this "laboratory." In this way he developed the concept of effective quantities, still valid today—"dosis sola facit venenum" ("the dose alone makes a thing poisonous").

Up until this time, folk medicine had been of great daily and practical significance. Persons who were in contact with nature and who observed animals were considered to be knowledgeable about the efficacy of different herbs, and this knowledge was recorded and passed on. Luther spoke strongly in his "Tischreden" (Colloquia Doct.): "I am astonished that God has set such high and noble medicine in dung; for we know from experience that pig manure stops the flow of blood. Horse manure can be used for pleurisy. Human dung heals wounds and hemorrhagic smallpox. Donkey manure is needed, among other things, for bloody dysentery, and cow dung with preserved

FIGURE 1 An apothecary's laboratory, 1560. Copper engraving from Ph.Galle after Stradamus around 1560, Sociopolitical Archive of Centre for Social Policy Research, University Bremen.

roses can be used for epilepsy in children." Luther was referring to a dubious increase in the wealth of available remedies that owed their existence to the so-called "Drecksapotheke" (filthy pharmacy) of Kristian Frantz Paullini (1696).

At that time universities were extended in the direction of natural sciences. While the earlier medical colleges (the first of which was set up in Salerno, by Naples) worked essentially from collections of books and translations, that is to say with the received wisdom of classical and Arab medicine, analytical endeavors now began. In 1533 the first chair of botany or materia medica was established in the medical faculty at the University of Padua. In 1609, Moritz, Landgrave of Hesse, appointed his personal physician, Johannes Hartmann (1568–1631), to the new professorship of "Chymiatrie" and thus established chemistry as an academic discipline, initially within the medical faculty. Such initiatives met with initial apprehension that chemical experiments and products could lead to health problems for local residents and neighbors, and corresponded to the unpleasant experiences that the common man (as opposed to the members of the noble ruling classes or the well-off middle classes) had undergone during medical treatment at the hands of contemporary physicians.

The results of scientific research and the development of more accurate methods led gradually, however, to an acceptance of scientific procedure, from which medicine also eventually profited. In 1628 William Harvey discovered the circulation of blood. In 1661 Robert Boyle (1627–1691) published his treatise "The Sceptical Chymist," in which he took a critical view of the accuracy and conclusiveness of many experiments. Georg Ernst Stahl (1659–1734), after making his own investigations, accused the alchemists of sloppy science and laid down precise methods of chemical research. The new developments took on visible form with the founding of the University of Halle in 1694, in which philosophy and natural sciences were brought together on the principle of freedom of thought and freedom of instruction.

A further thrust in the development of the sciences came through the discovery of America. While it is not clear whether the syphilis which raged through Europe in 1495, and which increased the importance of mercury, was brought by Columbus, some efficacious medical remedies were the result of his discovery—ipecac and cinchona, for example, were brought to Europe and promoted the demand and trade in pharmaceutical substances. In this period, which was ushered in by Paracelsus and which was characterized by the success of scientific modes of thought, toxicology was consolidated as a distinct field.

ACADEMIC QUALIFICATION AND THE PRACTICAL APPLICATION OF TOXICOLOGICAL KNOWLEDGE AT THE BEGINNING OF THE INDUSTRIAL AGE

The success of scientific modes of thought brought with it a demand for scientific–academic qualification. The trend in the Age of Enlightenment was that academic qualification should mean above all, training people to be socially and vocationally useful. In connection with this new direction in society towards, among other things, the advancement of trade and

industry, there followed the founding of university chairs. Under the influence of Justus von Liebig (1803–1873), the investigation of the causes of natural phenomena came to the foreground; chemistry in particular promoted the study of cause and effect.

In the 18th century, toxicology developed in the context of chemistry as a pure science and in the context of pharmacy as an applied science. It was taught, for example, by Georg Augustin Bertele (1767–1818), professor of pharmacy, toxicology, pharmacology, and the study of chemical formulae at the medical faculty in Landshut (which was later moved to Munich) and by his successor, Johannes Buchner (1783–1852). In the British colonies in North America, toxicology continued to be subsumed under the rubric of materia medica until 1765, when the first medical school was established at the College of Philadelphia.

At the level of practice and business, mechanical–economical chemistry developed. Under cameralism, the theory and practice of a well-ordered community at the end of the 18th century, manufacturing and factories took on great social significance. The profitable application of chemical principles to arts and crafts was at the center of this development. In the association between medicines produced for business profit and the health risks of particular trades, toxicology found a new emphasis. Henceforth it could be seen as an applied science, moving in step with the rapid industrial boom.

In occupational toxicology significant connections became apparent, which went along with very obvious risks and health impairment. This was so in mines, particularly in the case of carbon monoxide poisoning, but also in metal-processing industries (especially for lead and mercury). The Italian physician Bernardo Ramazzini (1633–1714) described in 1700 the diseases of craftsmen and artists so thoroughly and systematically that his work was translated into five languages and remained a textbook of occupational hygiene for a century.

Laurent Lavoisier (1743–1794) continued to concentrate on working empirically with substances and developed chemical research mainly on the basis of speculative experimental research (in the manner of Isaac Newton) and following the axiom of the conservation of matter. His "Traité Elémentaire de Chimie" appeared in 1789 and was to remain into the 19th century a shining example of what a chemistry textbook should be.

The industrial manufacture of Glauber's salt from rock salt and sulfuric acid used materials and their actions to a specific end. In the process, the technique

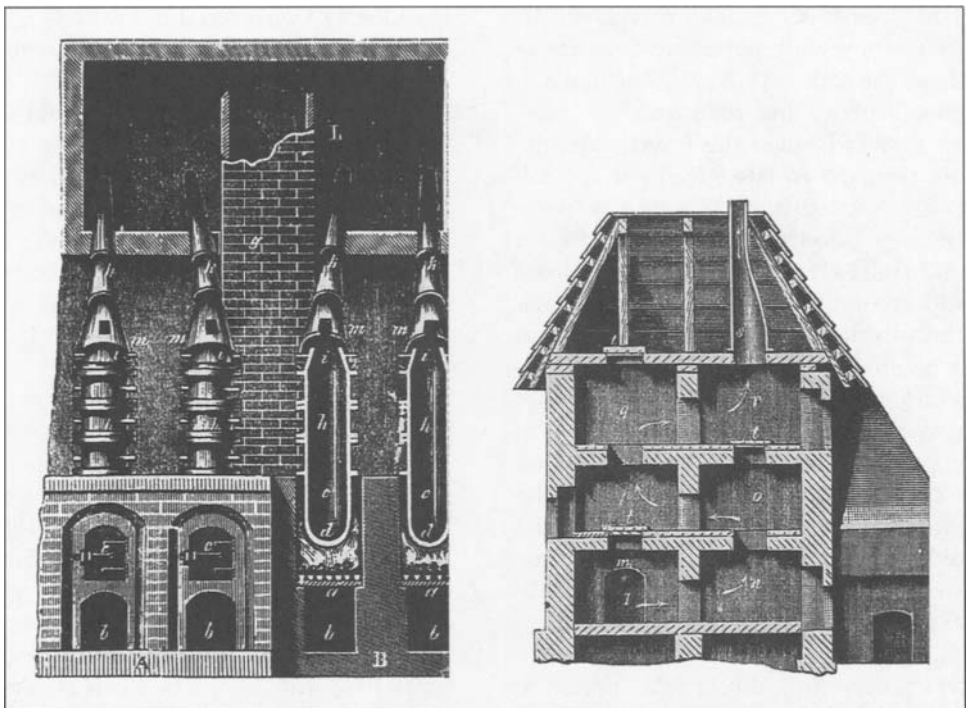


FIGURE 2 Poison processing towers and collection chambers for arsenious acid. From Hermann Scelenz (1911). Zur Geschichte der parmazeutish-chemischen Destilliergeräte. Hildescheim: Georg Olms, 1964, pp. 114, 115.

of making soda from sodium sulfate and coal was discovered. Along with sulfuric acid, soda was an all purpose material whose principal source was the soda lakes in Egypt. Now in 1790, with Nicolas Leblanc (1724–1806) and the process he developed, there began a chapter that is often described as the conquest of chemical discovery. The physician Leblanc, however, was a successful practical chemist for a short time only—only until his patrons were condemned to death and his soda factory was confiscated by a revolutionary tribunal. Leblanc committed suicide in the poorhouse.

In 1813 Mathieu J. B. Orfila summarized the results of specific toxic effects, added what was known to be possible in the way of precise detection of poison and in his "Traité de toxicologie générale" made a systematic compilation of the level of knowledge at the time.

CHEMISTRY AND THE CHEMICAL INDUSTRY IN THE NINETEENTH CENTURY

With technical production and the control of chemical activity, pharmacology became a new pure science. It kept pace with the most important advances that were being made in experimental physiology, especially under the influence of Johannes Müller (1801–1858) and his assistants. Unlike medicine, the methods of pharmacology were not dependent on investigations made at the sickbed. In 1802 the association between medical theory and pharmacology took an important step forward when the French apothecary Jean-Francois Derosne isolated the first crystalline "opium salt" (as a mixture of at least two alkaloids). The apothecary Friedrich Wilhelm Sertürner from Paderborn (1783–1841) created morphium from opium in 1804 and recognized in 1817 that the alkaline product is the physiologically active constituent of opium. At the beginning of the 19th century Parisian pharmacists (among them Pierre-Joseph Pelletier and Joseph-Bienaimé Caventou) discovered the substances chlorine, bromine, and iodine, and isolated quinine, caffeine, codeine, strychnine, and others. The state of learning was summarized by the physiologist François Magendie (1783–1855) who at the same time influenced research into the specific effects of substances in clinical medicine via his pupil Claude Bernard (1813–1878).

An important turning point, although it cannot be precisely dated, was the systematic pharmacological screening of synthetically produced molecules. In this way, pharmacologically effective substances could be made available in the form of preparations for therapeutic purposes, such as, for example, Serullass iodoform (1829) chloroform, obtained by both Soubeiran and Liebig at the same time in 1831.

As the importance of medicine and pharmacology grew, so did that of toxicology as an applied science. One of the influences on this development, along with pharmacy, continued to be the analytic impulse coming from forensic medicine.

Throughout history there has been an abundance of spectacular cases of murder through poison, from the female poisoner Locusta, who worked for Nero, to Teofania di Adamo, who was executed in Palermo in 1633 for murders using her Aqua Tofania (white arsenic boiled with lead and antimony), to Gesche Gottfried, who killed at least fifteen people and in 1831 was the last female criminal to be publicly beheaded in Bremen. During the 19th century most of the world-famous cases of systematic poisoning on a grand or a small scale involved the use of arsenic compounds, above all arsenious acid. As late as 1926 a nurse who had poisoned twelve people with arsenic was executed in Paris. J. Marsh took a significant step forward in 1832 when he invented an apparatus to show the level of arsenic present and was thus able to prove the use of the poison before an English court. The court, however, refused to recognise his evidence. It was not until 1840 that the Marsh apparatus given long-lasting recognition when it was used by the Parisian forensic medical expert Orfila in a sensational murder-by-poison trial in France.

Forensic medicine was initially taught, for example by Johann Peter Frank (1745–1821) in Pavia and Vienna, along with surgery. Around 1770 the Vienna School developed pharmacological and clinical experiments (Anton de Haen worked with opium, bearberry (*arctostaphylus uva ursi*), and cinchona). In 1804 it set up a separate professorship in a historical combination of state medicine and pharmacology, and in 1818 it established its own institute within the general hospital. The most important impetus came from Franz Cölestin Schneider (1812–1897) who succeeded, for example, in separating out arsenic in 1851 and who, like Pettenkoffer, worked towards combining chemistry and medicine. Out of this too grew the toxicological orientation of forensic medicine, for which Graz (Adolf Schauenstein 1827–1891) and Vienna (Ernst Ludwig 1842–1915, full professor of medical chemistry, 1874) were the main centers. In Zurich, in contrast, the disciplines were in competition, and as a consequence physiological chemistry was not offered until 1887 and Max Cloetta was not made associate professor of pharmacology until 1901. Cloetta represented a group of subjects, among them toxicology.

With Sertürner's isolation of pure morphine there began a new phase in cases of poisoning in which arsenic was supplanted by alkaloids, whose presence

in a corpse could not at first be proved. The number of possible poisons rose rapidly, while their detection during a postmortem examination grew ever more difficult. Carl Remigius Fresenius (1818–1897), on the other hand, was already demanding in 1844 that contradictory scientific presentations of evidence should be standardized and improved on. A spectacular trial took place in Belgium in 1851 (Count Bocarmé had poisoned his brother-in-law with nicotine) during which Jean Servais Stas succeeded in separating the alkaloids from body tissue by the removal of protein by alcohol and by ether extraction. Subsequently detection methods for countless alkaloids were developed. This "isolation of the poisonous substance" remained one of the cornerstones of forensic medicine and toxicology.

At the end of the 17th and beginning of the 18th century, a stricter scientific training for pharmacists was considered increasingly important. As early as 1779, the apothecary Johann Christian Wiegleb founded an institute for the teaching of pharmaceutics in Langensalza; more were established in the following years. The private institute of Johann Bartholomäus Trommsdorf in Erfurt (founded 1795) was famous; it served as a model for other institutes. In 1826 Liebig founded a private pharmaceutical–technological school in Gießen. Out of the larger pharmacies the first factories grew; Trommsdorf set up the first chemical–pharmaceutical factory in Erfurt in 1813, and his pupil Merck did the same in Darmstadt in 1827.

There was a close and fruitful connection between the academic science of the universities and commercial practice, illustrated, for example, by the relationship between Justus Liebig and Heinrich Emanuel Merck in the 1830s. In 1827 Merck began the factory production of morphine using the dried latex of opium poppy capsules. In the same year the Prussian government issued an edict that allowed doctors and pharmacists to obtain from factories medicines that were difficult to make. In England in 1844, William Brockedon compressed tablets out of a dry powder; shortly afterwards the pharmaceutical trade began to deal in gelatin capsules and coated tablets, followed by ampoules. There was a significant boom in the industrial production of medicines after 1870.

The development of pharmacology was one of the main reasons for an expansion in toxicology at university level in the second half of the 19th century. In 1847 Rudolf Buchheim founded his pharmacological laboratory in Dorpat, initially on a private basis, and introduced experimentation on animals. In the footsteps of Buchheim, who in 1860 set up the state institute for pharmacology, followed Rudolf Boehm, Hans Horst Meyer, and Rudolf Kobert. The center of emphasis shifted later with Oswald Schmiedeberg (1838–1921) to the newly established model university of Strasbourg. Out of private institutions, which were called "laboratories" until the 1870s, developed professorates and institutes in all the larger universities, in which toxicological approaches and concerns were also taken into account. S. Weir Mitchell, under the influence of Claude Bernard, carried out in 1858 a series of experimental studies on the toxicology of snake venoms and arrow poisons.

In the mid-1880s the systematic linking up of university scientific research with industrial laboratory work began. The intensity of research and specialization towards the end of the 19th century posed new problems for the integration of toxicology, particularly as a result of the growing independence of chemistry and pharmacology. As far as practical application was concerned, the scientific fields of pharmacy, toxicology, and food chemistry were integrated in certain pharmaceutical institutes, like the one in Marburg, which was developed by Robert Wilhelm Bunsen (1811–1899). In the USA, similar research institutes were founded based on the German example, beginning with Johns Hopkins University's 1876 Chair in Experimental Pharmacology.

In the latter third of the 19th century, a marked advance took place in mathematical and natural science teaching and research. The Chemistry Congress in Karlsruhe in 1860 stimulated an economic, scientific, and technical thrust forward that was scientifically strengthened by the benzol theory of 1865. The foundations of the theory, laid by August Kekulé (1829–1896), opened up the way for empirical, methodological searches for new dyestuffs. Certain scientific advances were influenced by the laboratories of the chemical industry. Hoechst, for example, appointed in 1883 the university professor A. Laubenheimer, who carried out bacteriological and serological research in close association with Emil A. Behring (1854–1917) and Paul Ehrlich (1854–1914). In 1876 the American Chemical Society was founded. In the wake of conflicts during the 1890s between more academically and more industrially oriented chemists, this body developed into four specialized organizations: the Division of Industrial and Engineering Chemistry (1907), the American Society for Pharmacology and Experimental Therapeutics (1908), the American Society for Agricultural and Food Chemistry (1908), and the American Oil Chemists Society (1909). In addition, an American Institute of Chemical Engineers was founded.

The toxicological implications of the revolutionary scientific and industrial developmental tendencies were already apparent at the end of the 19th century, but a blind eye was turned in the hope of manufacturing effective substances synthetically and systemati-

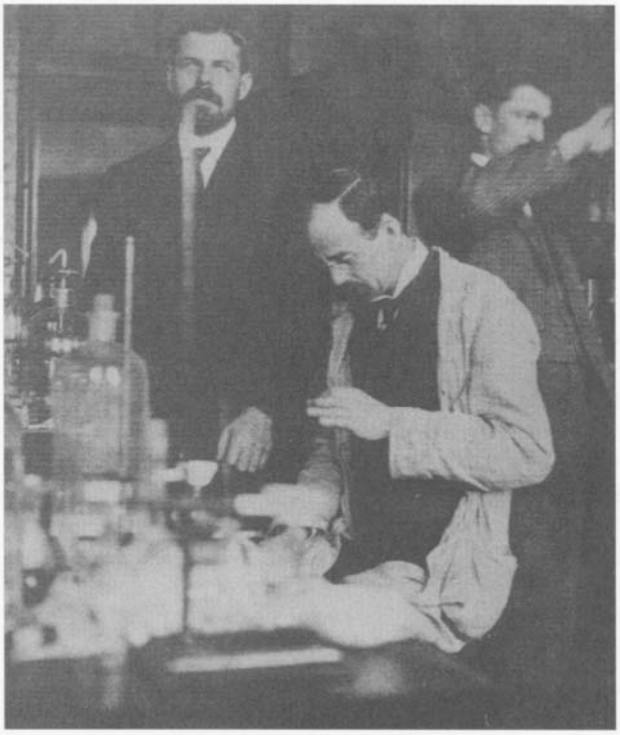


FIGURE 3 John Abel (seated) and colleagues at Johns Hopkins Laboratory, 1901.

cally. John J. Abel (1857–1938) studied in Germany under Carl Ludwig in Leipzig and Oswald Schmiedeberg in Strasbourg and became a founding father of pharmacology in the USA. Rudolf Kobert summarized the developments that had taken place in pharmacological knowledge at the end of the 19th century in two important textbooks, in which he emphasized the metabolic changes in the human organism through resorption, neutralization, oxidation, reduction, fission, conjugation, synthesis, secretion, and excretion. In his opinion, pharmacology was the fundamental science that was the impetus behind the applied branches of toxicology and pharmacotherapy.

SUBSTANCE EFFECTS—USES AND LIMITS

In the field of applied medical science, Charles Turner Thackrah (1795–1833) was the first to fashion an overview of studies in industrial hygiene, "The Effects of Arts, Trades and Professions on Health and Longevity" (1831), a pathbreaking study of worker health risks in England. A more useful and practical development arose from the association between chemistry, pharmacology, toxicology, and hygiene, as practiced by Max Pettenkofer (1818–1901). In 1852 the pharmacist and physician became a full professor of medical chemistry in Munich and in 1865 was appointed to the first professorship in experimental hygiene, also in Munich. Working in collaboration with Carl Voit, he began research on the exchange of gases between the organism and its environment.

Commercial developments augmented toxicological knowledge, which went beyond the knowledge of pharmacological agents and included all procedures. Hermann Eulenberg published the first summaries of results of experimental research on harmful and poisonous gases in 1865 and 1876, but had to admit that in practice one had to rely more on one's sense of smell than on chemistry. Ludwig Hirt (1844–1907) published in the 1870s the first modern standard work

on occupational medicine, which took as its starting point substances harmful to health.

The varying intensity and duration of cases of poisoning directed speculative attention to the difference between the substance itself and the affected individuals. In this regard Pettenkofer and his followers were particularly interested in the connection between the quantity of the substance and the body's processing of it. "Limits of toxicity," and "limits of tolerance," and the first limits for dangerous concentrations were calculated, among them the carbon dioxide concentration established in 1883 by Max Gruber. Karl Bernhard Lehmann (1858–1940) began in 1884 his research into the most important industrial gases. His institute in Würzburg (founded in 1887) became the foremost center of toxicological research. Up to the beginning of the Second World War he established the limits of concentration of over one hundred substances. Lehmann recognized, among other things, that in some gases relative habituation can occur, while others can cause increasingly severe reactions. In the USA, lists of harmful concentrations of gases had been established in 1912 and 1921. In 1927 the American Chemical Society published limits for 25 noxious gases concerning acute toxic affects and the "maximum concentration allowable for prolonged exposure." In 1946, the American Conference of Governmental Industrial Hygienists established 144 threshold-limit values for specific vapors, dusts, fumes, and mists.

At the turn of the century, pharmaceutical products became increasingly important in the development of the chemical industry. The manufacture of drugs shifted from the pharmacies to the factories and required precise analysis. As late as the beginning of the 19th century, the preferred method was experimentation on a small range of animals, the effects of which were as accurately described as possible, followed by a conclusion based on analogy with humans. Changes in behavior and secretions were noted. The tendency of the development was away from application according to symptoms and empirical discovery and toward systematically creating compounds and testing them.

The close association between commerce and medicine was henceforth influential in encouraging a connection, based on practical application, between toxicology and pharmaco-therapy. There was, for example, a collaboration between the firm of Hoechst and Paul Ehrlich, who was engaged in research into nontoxic arsenic compounds at the turn of the century. Chemotherapy, too, is linked with Ehrlich's name. His motto was: "Corpora non agunt, nisi fixata." In the 1880s Ehrlich was the first to establish the classification of leucocytes using a dye technique that is still in use today, and he was the founder of modern hematology. Ehrlich developed together with John N. Lanlay the drug-receptor theory, which in the 1930s was further developed by A. J. Clark.

In addition to the rapid developments in the field of pharmaceutics, important discoveries were made in the 19th century regarding the manufacture of indigo, sulfuric acid, ammonia, and chlorine, which influenced toxicology from the standpoint of occupational hygiene. From the 1890s onwards, chlorine was one of the keystones of organic synthesis. It made possible the use of chlorinated substances in the form of solvents, synthetic materials, wood preservatives, insecticides, herbicides, and so on. The decisive step for the future importance of pesticide toxicology was taken. Also decisive was the fact that the largescale availability of hydrogen chloride paved the way for the production of thermoplastic synthetic substances—polyvinyl chloride and polyvinyl acetate. The toxicity of these substances did not become significant until much later.

Many of the substances that were needed for industrial production were known to be dangerous, those used in the production of fuchsin, for example. At the Hoechst works in 1867, no less than 2.15 tons of 70% arsenious acid were processed. One could, however, according to Carl Duisberg, not only take measures to combat the dangers, but even remove them entirely, if the appropriate apparatus were used. Nevertheless, in 1903, more than 50 cases of aniline cancer were diagnosed among the workers at Badische Anilin- und Soda-Fabrik (BASF).

Thus the chemical industry at an early stage integrated medical knowledge and medical experts into the production to protect and select the workforce. Carl Knaps became the first factory physician at BASF in 1866; Wilhelm Grandhomme was appointed to the Hoechst works in 1874. These doctors assessed the toxicity of the raw materials, the factory buildings and installations, and the intermediate and the end products. The doctors of the large-scale chemical industries were of profound importance for the state of knowledge and the setting up of the areas of responsibility of the occupational hygiene authorities. In 1908, for example, the chemical industry set up the medical wing of BASF, out of which grew the "institute for occupational hygiene and toxicology." In collaboration with the employers' liability insurance association of the chemical industry, an occupational toxicology institute for research into and treatment of occupational disease was set up within the Ludwigshafen city hospital. In the context of the laws governing social insurance, in particular the laws regarding accident insurance (e.g., the U.S. Occupational Haz-

ards Act of 1889 and the British Workman's Compensation Act of 1906), toxicological questions had to be settled by expert opinion. The experiences of the First World War reinforced efforts of this kind, which found their expression in the inclusion of coverage for occupational disease—at first mainly occupational poisoning—in accident insurance. In England and the U.S., this authority lay predominately with government agencies (e.g., in the U.K., the Chief Inspector of Factories, created in 1896 and headed by Thomas Legge; and in the USA the Office of Industrial Hygiene and Sanitation, set up in 1919). The strengths of these agencies lay in field research and statistics rather than in laboratory research. In the Public Health Service, studies on industrial toxicology were carried out from 1919. In particular, Alice Hamilton conducted exemplary studies, for example, of the lead industry and of mercury poisoning in the felt hat industry. In 1919, she became the first female faculty member at Harvard University.

In Germany, toxicology experts, for example Louis Lewin, were caught in the sociopolitical struggle between labor and capital. Louis Lewin (1850–1929), assistant to Voit and Pettenkofer in Munich, qualified as a lecturer in pharmacology, toxicology, and hygiene in Berlin in 1881. His Jewish background was a particular hindrance to his career. He delivered his lectures in an apartment, which he had converted into an institute of pharmacology and toxicology. In addition to pharmacology and the effects and side effects of the newest drugs, he was engaged in toxicological work, which included occupational poisoning and how to prove a case of poisoning before a court. Lewin was one of the first teachers at an institute of higher learning to regularly take his students on visits to factories, in particular to chemical plants.

The International Association for the Statutory Protection of Workers, predecessor to the International Labor Office, acting on the advice of the occupational hygiene expert Theodor Sommerfeld, in particular, drew up a "List of Poisons" with the following aims: the obligation of doctors and hospitals to inform the authorities in the event of occupational poisoning, the preservation of the independence of company doctors, the obligatory registration of the production and use of commercial poisons, the institution of special morbidity identification papers for persons working with poisons, the promotion of toxicological research and teaching in medical science, the institution of special training for toxicologists, the expert supervision of work where poisons are used, and the institution of rules regarding hours of work for those employed in this field.

In these practical functions of occupational toxicology, findings were often made that were of far-reaching importance but that were only properly appreciated in other contexts. Thus the blister-like swellings in workers using fuchsin in an aniline factory had been described, but L. Rehn in 1895 traced these to the workers' exposure to aniline and fuchsin. For the first time the specific organotropic carcinogenic effects of chemical substances had appeared at a distance from the point at which they were caused. The first investigation of asbestos cases was conducted in 1906 by Montague and Murray, but did not bear policy consequences. The gap between initial scientific analyses and definite assessment in terms of occupational hygiene remained a practical difficulty for toxicology. K. B. Lehmann, for example, could give as early as 1898 a clinical picture of the disease chlorine-related acne among chemical workers involved in electrolysis. Later, after accidents in the 1950s, this was recognized as a symptom of massive exposure to dioxin.

Noise, vibration, air pollution (through smoke, soot, dust, foul smells, and toxic gases), ground pollution (through residue and wastewater) continued to be analyzed and assessed following Pettenkofer's tradition. The law "aerum corrumpere non licet," laid down by the *corpus iuris civilis* in the 17th century, was always relevant and topical in the period of industrialization, with its smoke from foundry chimneys, smog, and, above all, the toxic gases and vapors from factories.

In the field of food hygiene, too, we find important developments in the 19th century. A continual problem was, of course, alcohol. Earlier considered in many respects a remedy (to aid digestion), alcohol came to be seen from the 17th century onwards as a symbol of immoral behavior. Both public houses and pharmacies were now forbidden to sell brandy and beer on Sundays between and after the church services. Coffee, too, was originally considered a medical remedy, but its inclusion in the books of the apothecaries was short-lived. Soon it was being denounced as a new vice, was subject to taxes, and at the end of the 18th century was forbidden entirely. Looking at the matter from the point of view of economics, Frederick the Great considered it better to let the population re-accustom itself to beer. Food hygiene became increasingly important in Germany in the 1860s and 1870s, especially in the context of increasing amounts of "fake" foodstuffs, which became easier to detect with more effective methods in the 1830s. In 1879 the German Reich issued a decree concerning commerce in foodstuffs, alcohol, tobacco, stimulants, and utensils involved in food preparation. Initial attempts at

standardization in the U.S. occurred, for example, from 1901 at the Bureau of Standards, which made progress with the use of accurate methods of analysis. Under the auspices of the Pure Food and Drug Act, the Department of Agriculture began in 1906 to research toxicological questions. During the 1930s, above all in association with the mass production of canned goods, the Department of Agriculture's Food and Drug Administration carried out studies of the toxicity of substances used as food additives or that occurred as impurities in foods (e.g., insecticide residues). This was the beginning of broad-based campaigns, including among others the 1936 long-term project on the chronic toxicity of lead and arsenic and the Food, Drug, and Cosmetic Act of 1938.

Towards the end of the 19th century, research that was done into the problems of alcoholism and also into toxicomania took into account the aspect of addiction as well as that of depravation and underlined the connection between hygiene and efforts to improve public health.

Toxicology contributed to the recognition and minimization of several causes of danger to health present in foodstuffs. Food toxicology gained international recognition through scandals, when, for example, the hygienic conditions (or lack of them) in the meat-processing plants in Chicago became known. Another example is that of meat preservation. Until the beginning of the 20th century, meat was cured with a mixture of salt and saltpeter. J. Haldane established in 1901 that the desired red coloring of meat could be achieved by reducing the added nitrate to nitrite. The meat curers now began to use salt mixtures with pure nitrite. After curing salt with sodium nitrite came on the market in 1916, there were cases of poisoning, including a case of mass poisoning of 34 people, including a child who died, in Leipzig. Thereafter, the use of nitrite in curing was banned in Germany. The 1927 law relating to foodstuffs was drawn up on the basis of a list of prohibited additives, the criminally negligent use of which was punishable by law. The list was amended in 1958 to contain a "positive list" of all the specifically admissible substances.

There were difficulties with the analysis of those threats to health that have a long period of latency and complex causes. The carcinogenic effect of arsenic on humans was known because of its frequent use and was described as early as 1822. After the First World War, winegrowers combated the larvae of the grape moth with sprays containing arsenic. The risk of exposure was, however, limited to a period of a few days, and the toxicity of the related copper–arsenic compounds was known both to industry and to the authorities. In 1924 there were numerous reports from the Kaiserstuhl area of cases of arsenic poisoning and also of cancer. Franz Koelsch (1876–1940) pondered the question why workers in metallurgical plants and chemical factories, with greater exposure to these substances, almost never suffered from carcinomas or diseases of the liver. Severe chronic arsenic poisoning was also found among coopers, publicans, and winetasters. Only years later was the significance of the latency period recognized, when H. von Pein established in 1943 that 15 years or more can elapse between hyperkeratosis and the emergence of cancer, and that chronic arsenic poisoning can increase the organism's tendency to develop cancer. The assessment of arsenic poisoning in winegrowing remained a problem for toxicologists, however, because of the possibility of the misuse of alcohol.

Another famous example was the synthesis of dichlorodiphenyl-trichloroethane (DDT), which was carried out by the German chemist O. Zeidler as early as 1874, although the practical uses of the substance were not recognized until the Second World War when it was used as an insecticide. For this, the Swiss scientist Paul-Hermann Müller won the Nobel Prize for Physiology and Medicine in 1948. In 1972 special legislation was introduced to ban the use of the substance as an insecticide because of its unusual biological half-life. It had been realized that DDT (DDE) accumulates in the food chain. The scientific assessment of its practical application and use, missing in this case, became the central task of toxicology.

The seedtime of American industrial laboratories was the 1920s. The chemical industry, regarded as either a sort of glorified drugstore or as the producer of cheap and nasty substitutes, suddenly became the creator of new values. New close contacts between the chemical industry and the universities helped improve the heretofore somewhat suspect reputation of the drug industry. The chemical industry—and within it applied toxicology—developed in two spurts in the mid 20th century. First, the experience of the significance of Germany's poison gases and patented chemical technologies during the First World War, as well as the lessons learned from the analysis of and often dangerous application of petroleum in the booming interwar oil industry, boosted the importance of hygiene engineers and toxicologists in the U.S. Second, after the Second World War, in connection with the economic expansion, a boom in industrial hygiene and research into public health occurred, for instance, the Controlling Chemical Hazard reports conducted from 1945 to 1947 by the Bureau of Labor Studies. This development, in which German exiles

also participated, culminated in the Occupational Safety and Health Act of 1970.

From the middle of the 20th century onwards, experimental research on isolated organs and cell cultures was growing. This required specialized knowledge and made the practical generalization of the results difficult. Especially in regard to the statistical evaluation of the experiments, the quality of experimental planning and the collection and biometric processing of data have become an increasingly significant methodological problem. This is reflected today in the guidelines on "Good Laboratory Practice" (GLP).

A qualitatively new problem for toxicology arose with the discovery of radioactivity by H. Becquerel. Soon after this "sensational discovery," loss of hair and a decrease in the size of tumors were noticed among its biological effects. Once radium had been isolated by Marie and Pierre Curie, research into radiation and radioactive damage went hand in hand. The great attraction of X-rays and radioactivity initially masked the harmful effects of ionized radiation, to which Marie Curie herself fell victim. Safety precautions were called for as early as 1906, and in 1929 the International Commission on Radiological Protection was formed. The dosimeter for ionized radiation was originally developed for therapeutic purposes and later extended to cover all radiation effects. It was recognized early on that special limitations were necessary in the case of radiation damage to skin (Holthausen, 1936), to prevent genetic mutation in living creatures (Regaud and Dubreuil, 1908; Muller, 1927) and to avoid inducing cancer (Marie, 1910; Hesse 1911). In 1956 in the USA a threshold was set for genetic risk in the population (10 R in the first 30 years of life). The risks of radiation have become particularly significant in recent years, for sociopolitical reasons.

It must be borne in mind that there have been spectacular incidents in industrial production that have put great social pressure on toxicology, such as the mining disaster at Courrièrs in 1906, the chemical explosions in Ludwigshafen-Oppau in 1921 and in Buna in 1925, the smog disaster in the Maastaal in 1930, the mass poisoning due to a drug containing diethylene glycol in 1937, the fish poisoning in the Japan Kiushu due to industrial mercury residues dumped in the sea in 1953, the fetal poisoning caused by thalidomide from 1958–1961, the Santa Barbara oil spill in 1969, the trichlorphenol (dioxin) accident in Seveso in 1969, Love Canal in 1978, Three Mile Island in 1979, the poisonous gas explosion in the pesticides plant in Bhopal in 1984, the Institute of West Virginia in 1985, the admixture of diethylene glycol into Austrian wines in 1985, the introduction of 400 liters of herbicide into the canals and the Rhine river in Basel in 1986, and Tschernobyl in 1986. This pressure is not always conducive to good science.

Pharmacologists have always been interested in poisons, but in recent decades toxicology has become a separate discipline entirely, mainly on the basis of physiological and biochemical foundations. Still today, the subject of toxicology is found above all in the realm of biochemistry.

In forensic medicine, toxicological examinations are done in institutes of legal medicine and also in various bureaus of criminal investigation run by the German Bundesländer. There was reciprocal action among the demands of judicial presentation of evidence, analytic methods, and the quality of the institutes in which forensic toxicological work was being carried out.

In Zurich the medicolegal laboratory set up by Heinrich Zanggers became an official part of the university when he was appointed to an associate professorship there in 1906. In addition, there was the biochemical institute founded by Johann Bonifaz Flaschenträger (1894–1957).

Independent departments of toxicology were not established in universities until the 1960s. Environmental problems, above all, brought about the realization that the harmful effects of impurities on a healthy organism must be independently recognized, assessed, and prevented.

THE SCOPE OF TOXICOLOGY IN HISTORIC TERMS

"The question *why* is the mother of all science," said the philosopher Arthur Schopenhauer in 1813, and thus was formulated the guiding principle for the methodology of toxicology. The history of toxicology is the account of the endeavors made to answer that question qualitatively and quantitatively in a causal analytic way. It must be said, however, that the overlapping interpretations and evaluations of toxic effects are not part of the actual field of toxicology but represent indications and questions that toxicology must pass on to the fields of medicine, hygiene, or politics.

Oswald Schmiedeberg differentiated strictly between the "effects" and the "consequences" of substances. This differentiation characterizes the systematic orientation of toxicology. In its fundamental research into effects, toxicology was oriented towards medicine (physiology), chemistry and pharmacology. In the practical applied research into consequences,

toxicology was oriented towards pharmacy, forensic medicine, the chemical industry, and hygiene.

In a good historical survey written in 1925, Louis Lewin described toxicology in the widest sense as the study of life under the conditions that are known to cause illness. He believed that there were numerous variations in the conditions that cause illness and their corresponding physical malfunctions. The specific knowledge and experience that toxicology has developed can be used in the appraisal of the impact of these illnesses from a chemical, physical, and medical standpoint. Distinguished exponents classified the field of toxicological work as the common responsibility of civilized beings. As Lewin put it in 1922, "Anyone who is in the possession of an object which could be harmful to others is obliged to reduce or remove the possibility of that harm."

Bibliography

Ackerknecht. E. (1979). "Geschichte der Medizin." 4th ed. Enke, Stuttgart.

Amberger-Lahrmann, M., and Schmähl, D. (eds.) (1987). "Gifte: Geschichte der Toxikologie." Springer, Berlin.

Artelt, W. (1937). "Studien zur Geschichte der Begriffe 'Heilmittel' und 'Gift'." J. A. Barth, Leipzig.

Beyerlein, B. (1991). "Die Entwicklung der Pharmazie zur Hochschuldiziplin: Ein Beitrag zur Universitäts- und Sozialgeschichte." Wissenschaft, Verlagsgesellschaft, Stuttgart.

Chen, K. K. (ed.) (1969). "The American Society for Pharmacology and Experimental Therapeutics, Incorporated: The First Sixty Years, 1908–1969." ASPET, Bethesda, MD.

Cowen, D. L., and Helfand, W. H. (1990) "Pharmacy—An Illustrated History." Abrams, New York.

Doull, J. (1984). The past, present, and future of toxicology. *Pharmacol. Review* **36,** 15–18.

Doull, J., Klassen, C. D., and Amder, M. O. (1980). *The Basic Science of Poison* (J. Doull *et al.*, eds.) MacMillan, New York.

Earles, M. P. (1961). "Studies in the Development of Experimental Pharmacology in the Eighteenth and Early Nineteenth Centuries." London, University College.

Fischer-Homberger, E. (1983). "Medizin vor Gericht: Gerichtsmedizin von der Renaissance bis zur Aufklärung." Luchterhand, Bern.

Fye, W. B. (1987). "The Development of American Physiology. Scientific Medicine in the Nineteenth Century." Johns Hopkins University Press, Baltimore, MD.

Hamilton, A. (1925). "Industrial Poisons in the United States." MacMillan, New York.

Issekutz, B. (1971). "Die Geschichte der Arzneimittelforschung." Akadémiai Kiadó, Budapest.

Koelbing, H. M. (1985). "Die ärztliche Therapic: Grundzüge ihrer Geschichte." Darmstadt.

Lewin, L. (1922). "Die Gifte in der Weltgeschichte. Toxikologische, allgemeinverständliche Untersuchungen der historischen Quellen." J. Springer, Berlin.

de Malcissye, J. (1991). "Histoire du poison." F. Bourin, Paris.

Müller, R. K. (1986). "Dokumente zur Entwicklung der Toxikologie im 19. Jahrhundert." Akademie Verlagsgeselleschaft, Leipzig.

Parnham, M. J., and Bruinvels, J. (1983). *Discoveries in Pharmacology vol. 2.* Elsevier, Amsterdam, New York, Oxford.

Parascandola, J. (1990). "The Development of American Pharmacology." Johns Hopkins University Press, Baltimore, MD.

Perrine, D. M. (1996). "The Chemistry of Mind-Altering Drugs: History, Pharmacology and Cultural Context." American Chemical Society, Washington, D.C.

Sprung, W.-D. (1990). "Zur Geschichte von Pharmakologie und Toxikologie; Rudolf Kobert und seine Zeit; wissenschaftshistorische Beiträge zum Rudolf-Kobert-Symposium am 26, und 27, Oktober 1990 in Rostock." -University Abt. Wissenschaftspublizistik, Rostock.

Teleky, L. (1948) "History of Factory and Mine Hygiene." Columbia University Press, New York.

Tansey, T. (1948). The history of toxicology. The long and short of it. *Human & Experimental Toxicology,* **6,** 459–460.

Thackray, A. *et al.* (1985). "Chemistry in America 1876–1976." Reidel, Dordrecht.

Weindling, P. (ed.) (1985). "The Social History of Occupational Health." Helm, London.

Wetzel, W. (1991). "Naturwissenschaften und chemische Industrie in Deutschland: Voraussetzungen und Mechanismen ihres Aufstiegs im 19, Jahrhundert." F. Steiner, Stuttgart.

This Page Intentionally Left Blank

CHAPTER

2

Biostatistics in Toxicology

LUDWIG A. HOTHORN,* GERALD HAJIAN†

*University of Hannover, Hannover, Germany

†Schering–Plough Research Institute, Kenilworth, New Jersey

PRINCIPLES

This chapter covers the use of biostatistical methods in toxicology. Toxicology as a science is multifaceted; therefore, the methods presented are limited to regulatory toxicology studies. These studies are usually conducted according to national or international guidelines for the proof of safety for new drugs or chemicals. The problems are discussed from a biostatistical point of view. Selected simple statistical methods are described and demonstrated in small teaching examples, without need of computer programs.

At the end of a toxicological study, a decision about whether the new drug or chemical is harmless, harmful, or harmless up to a specified dose is made. Decision making requires a confirmatory type of statistical method given by the following criteria:

- A priori definition of an experimental design (frequently chosen according to guidelines).
- A priori definition of the necessary number of animals or sample sizes (frequently determined according to recommendations).
- A priori definition of the statistical test to be used and the upper limits of false positive (type I error α) and false negative (type II error β) decision rates.

The statistical methods can be classified into estimation (e.g., the estimation of LD_{50}) and testing (e.g., high dose versus negative control). The latter represents the emphasis of this chapter.

According to standard statistical techniques, the traditional null hypothesis of no difference in the effect between the treated and negative control group is tested.

Null hypothesis

H_0: effect of control group $\geq$ effect of dose group

Alternative hypothesis

H_A: effect of control group $<$ effect of dose group

Failing to reject the null hypothesis (i.e., when the p value is greater than 0.05), often leads to the conclusion that the evidence is in favor of safety (harmless). However, the most frequent testing theory is based on the falsification principle. Therefore, how the decision rule is set up in terms of the null and alternative hypotheses makes a big difference. Two approaches have to be distinguished: proof of hazard (null hypothesis of no difference) and proof of safety (alternative hypothesis of no relevant difference). In the first approach, the probability of erroneously concluding hazard (the producer's risk) will be controlled and in the second the probability of erroneously concluding safety (the consumer's risk) is limited. For drug-safety assessment, the consumer's risk (i.e., erroneously overlooking a toxic effect) should be controlled primarily. However, some difficulties still exist, for example, the a priori definition of the minimal relevant safety difference in a multiple endpoint case, and the approach represents a break from tradition. Therefore, the classical proof of hazard will be described here. However, a toxicological study must have a reasonable chance of detecting a relevant treatment effect. Consequently, the sample size should be based

on the type I and type II errors, and underlying variability, and the specific size of effects considered appropriate by the toxicologist. Moreover, in the case that the test failed to reject the null hypothesis, the type II error (or the power = 1 type II error) has to be reported.

Experimental design before the start of a study is important. Two aspects will be discussed here: sample size estimation, and choice of the many-to-one layout. However, in regulatory toxicology a minimum demand for sample size and number of dose groups is recommended by the guidelines. Another major objective of experimental design is dose selection. This is a very difficult task in real toxicological studies, and is not only dependent on statistical arguments. Therefore, this will not discussed here.

Sample Size Estimation

For a defined design and test the relationship between type I error α and type II error β is

$$\beta = f(n_j, \delta, \sigma, \alpha),$$

where δ denotes the minimal relevant difference, n_j the sample size per group, and σ^2 the variance. Type II error increases with smaller sample sizes n_j and/or increasing variance σ and/or smaller relevant difference δ and/or smaller α levels. As a simple example, sample-size estimation is demonstrated for the t test with assumed Gaussian distribution and equal sample sizes and variances:

$$n_j = (2\sigma^2/\delta^2)(z_{1-\alpha} + z_{1-\beta})^2$$

with $z_{1-\alpha}$ . . . $(1 - \alpha)$ quantile of normal distribution.

Sample-Size Estimation for a Subchronic Toxicity Study

For the serum variable, alkaline phosphatase (AP) in Wistar rat is known from historical controls to have a coefficient of variation (CV) equal to 36.5% (variance $\sigma^2 = 71$). Assuming a relevant safety difference $\delta = 4.2$ (i.e., 50% of the standard deviation) one-sided testing for an increase, sample sizes n_j are necessary for several levels of type II (0.05, 0.10, 0.25) and type I error (see Fig. 1).

In practice, such a blind sample-size estimation is often impossible because an unrealistically large sample size number would be calculated. A more pragmatic approach is the estimation of the type II error or the minimal detectable difference based on the sample sizes recommended by guidelines. Moreover, for each animal, multiple endpoints with different variances are measured. Based on the idea that a comparable type II error rate is important, the estimation of an endpoint-specific type I error rate is also possible. For the different endpoint (e.g., serum protein), an α level of 1.1% instead of 5% would have to be used to ensure the same β level.

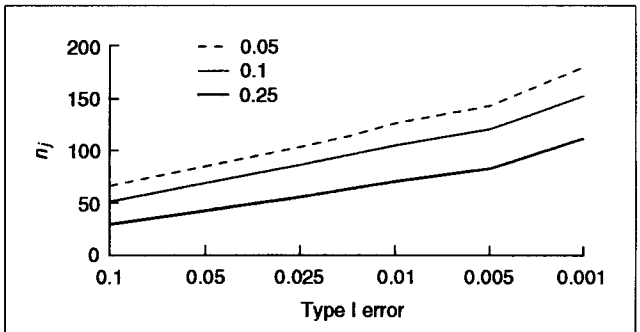


FIGURE 1 Relationship $n_j = f(\alpha, \beta, \sigma/\delta = 2, \text{one-sided})$ for the t-test.

Sometimes, the direction of the effect is known a priori; that is, only an increase of tumor rates represents the experimental question in a carcinogenicity study. Two minimize type II error, one-sided testing should be performed. If a two-sided test is used, a marked increase in sample size is necessary. In the above example, $n_{\text{one-sided}} = 32$, but $n_{\text{two-sided}} = 44$ is necessary.

In the many-to-one design, an increase in the sample size of the control will be used according to the statistical optimality rule $n_K = \sqrt{k}\, n_D$, where k denotes the number of dose groups. However, this rule holds true only in the case of variance homogeneity (which is questionable in toxicological studies, a priori) and the exact $\sqrt{k}$ rule is not necessary. For the typical three- or four-dose experiment, simply doubling the sample size is sufficient. Thus, the control sample size can be divided into two independent randomized controls, possibly stratified on pretest conditions (e.g., body weights). In the case of dichotomous data, such as mortality or tumor rates, the sample size n_i also depends on the spontaneous rate of the controls p_C (see Fig. 2).

Therefore, detection sensitivity increases with lower spontaneous rates. The minimum sensitivity occurs when is $p_C = 0.5$. In a carcinogenicity study based on a balanced false-positive–false-negative ratio, animal strains with the spontaneous rate of the target tumor greater than zero and smaller than about 10% should be selected. Unfortunately, the target tumor seldom can be defined a priori, because numerous tumors with rather different spontaneous rates must be taken into account. Therefore, an animal

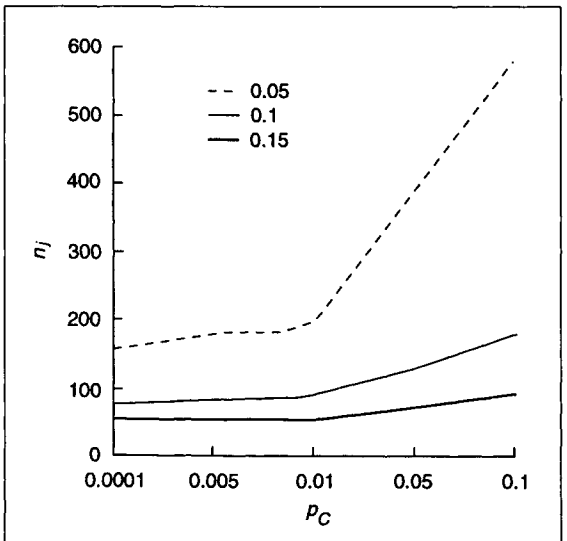


FIGURE 2 Sample-size estimation for the asymptotic test for comparing two proportions (one-sided, $\alpha = 0.05$, $\beta = 0.20$) as a function of spontaneous rate p_C and $\Delta p(0.05; 0.01; 0.15)$.

strain with an average small spontaneous rate across all tumor types should be selected.

Experimental Design

The many-to-one design *treatments versus negative control* (i.e., $k + 1$ experimental groups) is applied to all toxicological studies, with the exception of acute toxicity studies. Usually, this one-way layout is included in a factorial design with the factors of litter, sex, and time. Assuming an interaction between these factors in advance, an independent analysis is done for a dose effect by each factor separately. Therefore, only methods for the one-way design are described here and the following types can be distinguished:

- Unrestricted design: $C^-, T_1, \ldots, T_k$, with T_j denoting different treatments (e.g., different substances, application forms, or combinations of substances).
- Dose–response design: $C^-, D_1, \ldots, D_k$, with D_j denoting different doses; $C^- = 0 < D_1 < \cdots < D_k$ (k not too large, i.e., $\in \{2, 3, 4\}$).
- Extended design including a positive control: C^-, $D_1, \ldots, D_k, C^+$.
- Complex design: $C^-, C^0, S_A (D_1, \ldots, D_k)$, $S_B (D_1, \ldots, D_k), C^+$, where C^0 denotes an empty control (without any administration), and $S_{A,B}$ denote several substances.

For the first type, comparisons of treatments versus negative control (many-to-one procedures) are used. For the second type, the following can be distinguished:

- Dose–response analysis based on selected complex models (e.g., single hit model) and testing selected model parameters. Unfortunately, there is no possibility for the separation of model and lack-of-fit. According to the slogan "all models are wrong, some are helpful," independent model selection for the typical three- or four-dose design is difficult and the routine use of this approach cannot be recommended.
- Dose–response analysis with comparison procedures for ordered alternatives. A priori assumption that the rejection of the null hypothesis will be according to an increasing trend only:

$$H_0: \quad \mu_C = \mu_{D_1} = \mu_{D_2} = \cdots = \mu_{D_k}$$

$$H_A: \quad \mu_C \leq \mu_{D_1} \leq \mu_{D_2} \leq \cdots \leq \mu_{D_k}$$
$$\text{with at least } \mu_C < \mu_{D_k}$$

Weaker assumptions are used (monotonic increasing, without knowledge of the kind of shape); therefore, this method is more robust.

Comparison Procedures "Control versus k Treatments or Doses"

In several publications, numerous statistical approaches for many-to-one designs in toxicology were described. Two concepts should be distinguished:

- Pairwise tests (t tests, u tests) "control versus group j," each independent at level α. Here, the principle of an experimental type I error rate is violated and no order information is used. Independent of the number of groups, a minimum type II error rate is guaranteed.
- Multiple comparison procedures guarantee an experimental type I error rate, but depending on the number of groups, the type II error rate will increase. If a priori monotonicity of the dose–response can be assumed, either procedures for order restriction or trend tests within the closure test principle should be used to decrease the type II error.

However, the following approaches should be avoided:

- Pairwise tests "each against each other" using t or u tests, because the experimental α level will be strongly violated.

TABLE 1 Recommendations for Pairwise Tests

Endpoint	Example	Two-sample test
Parametric, homogeneous variances	Hemoglobin	t test
Parametric, heterogeneous variances	Erythrocytes	Welch t test
Nonparametric, continuous	ASAT	u test, asymptotic version
Nonparametric, discrete with ties	Number of micronuclei	u test, permutative version
Dichotomous	Mortality rate	Fisher's or Barnard's test

- All-pair comparisons procedures (e.g., the well-known Tukey procedure) because the many-to-one structure is ignored and type II errors increase dramatically.
- "Omnibus" global tests (e.g., Kruskal and Wallis (1952) test or F test) because the rejection of the null hypothesis only shows heterogeneity among the groups in an unknown structure.
- Test hierarchies consisting of test on distribution, omnibus test, and procedures because, from the viewpoint of type II error, they are not free of contradiction.

The choice depends on the type of endpoint, which is shown for pairwise tests in Table 1 and many-to-one procedures in Table 2 (the tests and procedures used are explained in the examples).

Stepwise modifications of most procedures are now available. They guarantee smaller type II errors. However, no simultaneous confidence intervals are generally available. Until now, only "mean value statistics" (e.g., t test or Dunnett's procedure) have been considered in this paper. Frequently in toxicological studies, a responder–nonresponder behavior can be observed. Tests on means are based on the principle of 100% responders. For cases where responders are less than 100%, tests for Lehmann alternatives can be used. In this case, we assume that the treatment consists of a proportion of p nonresponders (i.e., the same behavior as control group animals) and a proportion $(1 - p)$ of pathologically reacting animals. Such a mixing distribution can be detected by larger variances, which is exactly what is frequently observed in toxicology studies: increasing variances with increasing effects (see Conover and Salsburg, 1988).

Robustness

The tests and procedures should be robust against violation of their assumptions, that is, no increase of type I or type II error rates due to real data problems. Real data are characterized by violations of the Gaussian distribution and/or the variance homogeneity assumption. Unfortunately, in small-sample-size studies, a reliable test for a distribution is impossible. Either such a test can be performed on historical controls or a nonnormal distribution is assumed a priori, leading to nonparametric methods. However, numerous investigations have revealed that under different distribution types usually observed in toxicology, the parametric tests and procedures are relatively robust. In contrast, both parametric and nonparametric methods are very sensitive to heterogenous variances, especially when combined with unbalancedness (high variances and unequal small sample sizes). Variance

TABLE 2 Recommendations for Many-to-One Procedures

Endpoint	Without restriction of the alternative hypothesis	Ordered alternative hypothesis
Parametric	Dunnett (1955) procedure	A priori ordered procedure based on t tests or contrasts
Nonparametric	Steel (1959) procedure	A priori ordered procedure based on u tests or Jonckheere (1954) trend tests
Dichotomous	Passing (1984) procedure	A priori ordered procedure based on Fisher tests or Armitage (1955) trend tests

heterogeneity can be simply tested on the concurrent study (e.g., using the F test or the Levene 1960 test). The Levene test has the added benefit of being robust under skewed distributions and hence is recommended. If the variance homogeneity is markedly violated, procedures based on common mean-square error should not be used. Simply α-adjusted Welch t tests can be used instead.

Reporting of "Significances"

The results of significance tests or procedures can be reported as:

- Significant or not significant as a dichotomous decision.
- Symbols (e.g., * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$).
- p values (the smallest possible empirical type I error rate).
- Confidence intervals for the differences control versus treatment j.

While p values are the shortest representation and are now up-to-date, parametric confidence intervals $[x_{\text{lower}}; x_{\text{upper}}]$ contain the most information. The distance from zero is a measure of the magnitude of violation of the null hypothesis; the width is a measure of variance, sample size, dimension k, and type I error α. Confidence intervals are scale variant (endpoint-specific interpretation is easy) and can be used in the safety approach (a posteriori defined minimal relevant difference δ). For Dunnett's procedure (see the section on Chronic Toxicity) the two-sided $(1 - \alpha)$-confidence interval is

$$[\bar{x}_j - \bar{x}_C - d_{k,\nu,\rho_{Cj},1-\alpha}\sqrt{s^2(1/n_j + 1/n_C)};$$

$$\bar{x}_j - \bar{x}_C + d_{k,\nu,\rho_{Cj},1-\alpha}\sqrt{s^2(1/n_j + 1/n_C)}],$$

where $d_{k,\nu,\rho_{Cj},1-\alpha}$ denotes the $(1 - \alpha)$-quantile of the multivariate t distribution with correlation ρ_{Cj}. Selected quantiles are given in Table 3 (calculated using the PROBMC function in SAS).

Descriptive Statistics

Not only should the results of significance tests or procedures (p values or confidence intervals) be reported, but also group-specific location, scale measures, and sample sizes. Either parametric statistics (mean and standard deviation) or nonparametric (median and percentiles) should be chosen.

ACUTE TOXICITY

The objective of an acute toxicity study is the estimation of the toxicity of a compound after the single administration of increasing doses to estimate, among others, the lethal dose for 50% of the animal population: LD_{50}. A randomized one-way design (independent for each sex) with the factor dose at the levels $D_1 < D_2 < \cdots < D_k$ will be used. The endpoint is the number of dead animals r_j in group j after a predefined observation time (e.g., 48 h) in relation to the number of animals at risk (e.g., group sample size n_j). This mortality rate $p_j = r_j/n_j$ represents a dichotomous variable and the statistical problem consists in the estimation of the dose–response relationship with the following steps:

- Transformation of p_j into an approximate normally distributed variable y_j.
- Logarithmic transformation of the dose D_j into $x_j = \log D_j$.
- Parameter estimation in the linear model $y_j = a + b\,x_j$.
- Estimation of the LD_{50} by inverse transformation:

TABLE 3 Selected 95%-Quantiles of the Multivariate t Distribution and Normal Distribution (equal correlation coefficients $1/\sqrt{2}$; n_j = const.)

	One-sided		Two-sided	
Degree of freedom ν	$k = 2$	$k = 3$	$k = 2$	$k = 3$
10	2.15	2.34	2.57	2.76
15	2.07	2.24	2.44	2.61
20	2.03	2.19	2.38	2.54
30	1.99	2.32	2.15	2.47
60	1.95	2.10	2.27	2.41
120	1.93	2.08	2.24	2.38
∞ (multivariate normal distribution)	1.92	2.09	2.21	2.35

$p_j \longrightarrow 0.5\ (=50\%) \longrightarrow y_j = 0 \longrightarrow x_j \longrightarrow \mathrm{LD}_{50}$.

- Estimation of the two-sided $(1 - \alpha)$ confidence interval for LD_{50}.

In the statistical literature, the problem of different transformations of dichotomous variables is discussed. However, for the estimation of a central parameter like LD_{50}, the choice is not critical because well-known transformation methods, such as

Probit: $\mathrm{probit}(r_j/n_j) = \Phi^{-1}(r_j/n_j)$
with Φ^{-1} inverse of normal distribution,

Logit: $\mathrm{logit}(r_j/n_j) = \ln[(r_j/n_j)/1 - (r_j/n_j)]$, and

Angular: $\mathrm{angular}(r_j/n_j) = \arcsin\sqrt{r_j/n_j}$,

differ markedly only for $p_j < 0.1$ and $p_j > 0.9$. Because of the simple numerical property of the logit transformation, we will describe its use. Parameters can be estimated by the iterative maximum likelihood method, the method of moments, and the weighted least-square method. The last is simple and will be demonstrated here. A correction of the transformation

$$y_j = \ln\left(\frac{r_j + n_j/40}{n_j - r_j + n_j/40}\right)$$

is necessary because $p_j = 0$ and $p_j = 1$ are possible. Because $\mathrm{var}(y_j) \approx p_j(1 - p_j)/n_j$, the variance homogeneity assumption is violated and hence the weighted least-squares estimation with the weights $w_j = n_j(p_j + n_j)(1 - p_j^2)$ should be used with the asymptotic estimators

$$a = \bar{y} - b\bar{x}$$

$$\text{with } \bar{y} = \left(\sum_{j=1}^{k} w_j y_j\right) \Big/ \left(\sum_{j=1}^{k} w_j\right)$$

$$\text{and } \bar{x} = \left(\sum_{j=1}^{k} w_j y_j\right) \Big/ \left(\sum_{j=1}^{k} w_j\right)$$

where

$$b = \left(\sum_{j=1}^{k} w_j(y_j - \bar{y})(x_j - \bar{x})\right) \Big/ \left(\sum_{j=1}^{k} w_j(x_j - \bar{x})^2\right).$$

The estimation for the LD_{50} is simply $\mathrm{LD}_{50} = \exp(-a/b)$. The two-sided $(1 - \alpha)$ confidence interval of the LD_{50} is the *quotient of two variables a/b* and is given by Fieller's theorem:

$$\mathrm{CI}(\mathrm{LD}_{50}) = \exp\left(-\frac{a}{b} \pm z_{1-\alpha/2}\sqrt{\mathrm{Var}(\mathrm{LD}_{50})}\right)$$

with

$$\mathrm{Var}(\mathrm{LD}_{50}) = \left[\frac{1}{\sum_{j=1}^{k} w_j} + \mathrm{LD}_{50} - \frac{\bar{x}^2}{\sum_{j=1}^{k} w_j(x_j - \bar{x})^2}\right] / b^2$$

Numerous computer programs are available to produce these computations. Other relevant practical methods are nonparametric and sequential approaches for the estimation of approximate toxicity limits that reduce the number of experimental animals needed.

CHRONIC TOXICITY

Introduction

The objective of a repeated toxicity study is to estimate the toxicity and side effects of a new substance after repeated administration of increasing doses. Multiple comparisons of all the endpoints in treated groups versus the negative control group C^- determines the toxicological effects. The duration of administration varies from 10 days to 2 years, depending on the objective in relation to clinical trials and international regulatory recommendations.

The usual design is the completely randomized three-factorial layout:

SUBSTANCE × SEX × TIME
(where × denotes cross classification)

As levels for the factor substance, either treatments T_j (e.g., different substances, different application routes) or dose groups D_j are used in a many-to-one design:

$$C^-, T_1, \ldots, T_k \quad \text{or} \quad C^-, D_1, \ldots, D_k$$

As levels of the factor time, the measurement time points $t_1, t_2, \ldots, t_T$ are used; these are frequently nonequidistant. A typical characteristic of chronic toxicity studies is the simultaneous estimation of different measurements and findings for each individual animal (e.g., body weight, hematology, clinical chemistry parameters, urine analysis, organ weights, and clinical and histopathological findings). Although a multivariate analysis is possible, the state of the art is the univariate analysis done separately for each endpoint.

The primary problem for statistical analysis is the proof of substance effect by comparing the effect of the dose or treatment group with that of the negative control group. Although the three-way layout can be analyzed by standard analysis of variance methods, including the interactions dose × time and dose ×

TABLE 4 Repeated Toxicity Study with Number of Animals Dead after 6-Month Duration

	0	10 mg/kg	50 mg/kg	100 mg/kg	Σ
Number of animals dead, r_j	0	1	6	8	$r. = 19$
Number of animals at risk, n_j	40	20	20	20	$n. = 100$
Mortality rate, p_j	0.10	0.05	0.30	0.40	

sex, many-to-one comparisons frequently will be performed separately for each time point and sex.

Analysis of Mortality

Two types of mortality data can be distinguished: one with time dependence and the other without time dependence. The first case typically occurs in long-term studies (i.e., comparison of survival functions), while the second one occurs in short-term studies. In the last case, both the number of dead animals at the end of the study r_j and the number of animals at risk n_j are recorded (Table 4). The comparison of the proportion of the negative control $p_C = r_C/n_C$ with the proportion in the treatment groups p_j can be performed with the many-to-one procedure. Assuming a monotone dose–response relationship, the trend test can be used within a closure test principle according to Marcus *et al.* (1976).

Many-to-One Comparison Procedure for Dichotomous Variables (Passing, 1984)

$$dp_j = (p_j - p_C)/\sqrt{r.(1 - r./n.)(1/n_j + 1/n_C)(n./(n. - 1))} \quad \text{for each } j \in \{1, \ldots, k\}$$

with

$$p_j = r_j/n_j \quad r. = \sum_{j=C}^{k} r_j \text{ and } n. = \sum_{j=C}^{k} n_j$$

TABLE 5 Selected 95%-Quantiles of the t Distribution and Normal Distribution

ν	One-sided	Two-sided
12	1.782	2.179
14	1.761	2.145
16	1.746	2.120
18	1.734	2.101
20	1.725	2.086
40	1.684	2.021
60	1.671	2.000
∞ (normal distribution)	1.6449	1.9600

and

$$dp_{10\,mg/kg} = (0.05 - 0.10)/\sqrt{(19(1 - 19/100)(1/20 + 1/40)(100/100 - 1)} = -0.46$$

$$dp_{50\,mg/kg} = 1.85$$

and

$$dp_{100\,mg/kg} = 2.78$$

The null hypothesis of equality of the mortality rates of the treatment groups in comparison to the negative control will be rejected if the test statistic is larger than the quantile of the multivariate t distribution $t_{k,\text{one/two-sided},1-\alpha}$ (see Table 3). Only $dp_{100\,mg/kg} = 2.78$ is larger than the quantile $t_{k=3,\text{one-sided},1-\alpha=0.95} = 2.09$, and hence we can conclude that there is an increase of mortality in the 100 mg/kg group.

Trend Test For Dichotomous Variables (Armitage, 1955)

$$ta = \frac{\sum_{j=C}^{k} D_j(r_j - (n_j r./n.)}{\sqrt{r.(n. - r.)/(n.(n. - 1) \times \sum_{j=C}^{k} n_j\left(D_j - \left(\sum_{j=C}^{k} n_j D_j/n.\right)\right)^2}}$$

$$ta = \frac{0 + 10(1 - 20 \times 19/100) + 50(6 - 20 \times 19/100) + 100(8 - 20 \times 19/100)}{\sqrt{(19(100 - 19)/100 \times 99)(40(0 - 32)^2 + 20(10 + 32)^2 + 20(50 - 32)^2 + 20(100 - 32)^2}}$$

$$= 3.29.$$

The null hypothesis of equality of the mortality rates will be rejected if the test statistic is larger than the quantile of the normal distribution $z_{1-\alpha}$ (see Table 5).

TABLE 6 Analysis of Body-Weight Data by the *t* Test and AUC Approach

	Body weight/g	
t	Control	Dose
0	203, 206, 201, 196, 201	177, 183, 209, 189, 185
1	239, 243, 244, 239, 251	205, 209, 221, 191, 207
4	302, 291, 287, 291, 315	233, 251, 268, 227, 251
8	388, 371, 383, 371, 413	277, 322, 357, 270, 329
13	411, 391, 401, 399, 390	355, 299, 391, 311, 351

Because $ta = 3.29$ is larger than the quantile $z_{0.95} = 1.6449$, the null hypothesis can be rejected and a dose trend in mortality can be assumed. The closure principle can be used here to detect the minimum effective dose simply by conditional step-down testing of k, $k-1, k-2, \ldots, 1$ dimensional trend tests, each at level α (if the null hypothesis in the jth dimension can be rejected, the $(j-1)$ dimensional trend test can be tested; otherwise, the procedure stops and the minimum effective dose is j). In the example for the four-dimensional trend test (global test) the null hypothesis can be rejected. The test statistic for the three-dimensional trend test is 2.27 and again larger than 1.6449. Therefore, the two-dimensional trend test can be conducted and because this test value is 0.65, the minimum effective dose is 50 mg/kg.

Analysis of Body Weight

Two types of body weight data can be distinguished: (1) only a few measurements (e.g., weekly in a 4-week study) and (2) growth curves based on numerous measurements (from a 2-y carcinogenicity study). In the first case, individual differences from the baseline are calculated and separate tests at each time point are appropriate. This approach could also be used in the second case; however, the large number of tests will increase the global false-positive rate. Body weight-versus-time relationships are important in long-term studies because this noninvasive simple method has a large predictive value for potential toxic effects. The problem is the comparison of the growth curve of the negative control versus those of the treatment or dose groups under the conditions of nonequidistant measurements and, dose- and time-dependent drop out due to mortality. Three simple approaches will be recommended here:

- Fitting a nonlinear growth curve model with the smallest possible number of parameters and testing between selected parameters for treatment effects.
- Assuming the repeated measures are multivariate vectors and using multiple endpoint test in its stabilized version.
- Transformation into an univariate individual variable in the sense of an integral measure of the growth curve using the area-under-the-curve (AUC) technique for the global time interval $(0,T)$ or selected time intervals (t_1, t_2). AUC can be computed simply by trapezoidal rule.

In Table 6, the t test and the AUC approach are demonstrated for a small number of real data. Only a part of the raw data from a 90-day study on female rats is presented. The test is one-sided for a decrease.

t Test for Differences to Baseline

Differences in week 4;

Control: 99, 85, 86, 95, 114

Dose: 56, 68, 59, 39, 66

The p value of the t test is 0.0004 and the null hypothesis can be rejected.

AUC Approach

Individual AUC estimation using trapezoidal rule between t_0 and t_{end} (0th and 13th weeks):

$$\text{AUC} \approx (m_0(t_1 - t_0) + m_1(t_2 - t_0) + m_2(t_3 - t_1) + \cdots + m_{end}(t_{end} - t_{end-1})/2.$$

For the first animal in the control, the AUC is

$$(203*1 + 239*4 + 302*7 + 388*9 + 411*5)/2$$

For all other animals, the AUCs are:

Control: 4410, 4254.5, 4319, 4261.5, 4538.5

Dose: 3448, 3584.5, 4068.5, 3263.5, 3743

TABLE 7 Subchronic Toxicity Study on Wistar Rat, Endpoint Alkaline Phosphatase

Group	n_j	Alkaline phosphatase	Mean	SD
Control	12	17.3, 23.0, 22.5, 10.0, 21.6, 19.7, 17.7, 20.0, 17.9, 20.9, 24.3, 19.9	20.317	2.216
D_{low}	12	23.9, 19.2, 24.1, 19.1, 17.4, 21.9, 24.1, 24.0, 20.7, 21.9, 23.1, 16.9	21.358	2.660
D_{high}	10	24.2, 31.3, 30.7, 17.4, 27.9, 25.7, 18.1, 21.9, 25.9, 19.8	24.290	4.939

Now the mean values and standard deviations from these transformed values are calculated. Because the variance is heterogeneous, the Welch t test was used.

$$t = \frac{\bar{x}_D - \bar{x}_C}{\sqrt{s_D^2/n_D + s_C^2/n_C}}$$

with

$$\nu = \frac{(s_D^2/n_D + s_C^2/n_C)^2}{(s_D^2/n_D)^2/(n_D - 1) + (s_C^2/n_C)^2/(n_C - 1)}$$

With the modified degrees of freedom $\nu = 5$, a p value of 0.0021 results; that is, the integral of the body weight–time relationship from the beginning to the end is significantly reduced.

Analysis of Continuous Variables

The choice of the procedure depends on the type of design, type of endpoint, and the possibility of order restriction. In Table 7, the endpoint alkaline phosphatase was measured in a many-to-one design.

Parametric Many-to-One Procedure (Dunnett, 1955)

$$d_j = \frac{\bar{x}_j - \bar{x}_C}{\sqrt{\text{MSE}(1/n_j + 1/n_C)}} \quad \text{for both } j \text{ (i.e., } D_{\text{low}}, D_{\text{high}})$$

$$\text{with MSE} = \frac{\sum_{j=C}^{k}\sum_{i=1}^{n_j} x_{ij}^2 - \sum_{j=C}^{k} x_{j\cdot}^2/n_j}{\sum_{j=C}^{k}(n_j - 1)}$$

$$\text{and } \nu = \sum_{j=K}^{k}(n_j - 1)$$

$$\text{and } \rho_{Cj} = 1/\sqrt{n_j/(n_C + n_j)}$$

TABLE 8 Selected Values f for Nonparametric, Two-Sided $1 - \alpha = 95\%$ Tolerance Intervals for $p = 90$ or 95% of the Population[a]

Sample size of historical controls	Value f	
	$p = 90\%$	$p = 95\%$
30	1	1
50	2	1
100	5	2
200	13	5
300	22	9
400	30	13
500	39	17

[a]According to Hahn and Meeker, "Statistical Intervals," Copyright © 1991. Adapted by permission of John Wiley & Sons, Inc.

The null hypothesis for group j is rejected if $d_j > d_{k,\nu,\text{one/two-sided},\rho_{Cj},1-\alpha}$, where $d_{k,\nu,\text{one/two-sided},\rho_{Cj},1-\alpha}$ denotes the quantile of the multivariate t distribution with degree of freedom ν and correlation coefficients ρ_{Cj} (see Table 3). The values of the statistics are

$$d_{D_{\text{high}}} = \frac{(24.29 - 20.317)}{\sqrt{11.33(1/12 + 1/10)}} = 2.74$$

and

$$d_{D_{\text{low}}} = \frac{(21.358 - 20.317)}{\sqrt{11.33(1/12 + 1/12)}} = 0.756 \quad (\text{with MSE} = 11.33)$$

Because $d_{k=2,\nu=31,\text{one-sided},\rho=.707;0.74,1-\alpha=0.95} = 1.983$ (estimated by the PROBMC function in SAS; the approximate table value is 1.99; see Table 3), the high dose is increased with respect to control. The stepwise procedure is as follows: (1) first order the test statistics d_j, (2) then compare the largest test statistics with the quantile $d_{k,\nu,\text{one/two-sided},\rho_{Cj},1-\alpha}$, (3) if and only if this null hypothesis is rejected, compare the next largest test statistics with the quantile $d_{k-1,\nu,\text{one/two-sided},\rho_{Cj},1-\alpha}$ (i.e., compare with the k, $(k - 1)$, $(k - 2)$, . . . , one-dimensional quantile instead of only the k-dimensional quantities). This approach decreases the type II error.

Nonparametric Many-to-One Procedure (Steel, 1959)
***Pairwise Ranking {C,D$_j$}*, Asymptotic Version:**

$$ds_j = \frac{\sum_{i=1}^{n_j} R_{ij} - n_j(n_j + n_C + 1)/2}{\sqrt{n_C n_j(n_j + n_C + 1)/12}}$$

with R_{ij} . . . pairwise ranking of C and D_j

The null hypothesis is rejected for group j if $ds_j > d_{k,\nu=\infty,\text{one/two-sided},\rho_{Cj},1-\alpha}$ (see Table 3). The pairwise ranking works for the low dose with respect to the control group as follows: (1) rank the pooled sample from 1 to $(n_C + n_j)$ (mid-ranks for equal values):

C: 3.0, 18.0, 17.0, 1.0, 14.0, 9.0, 5.0, 11.0, 6.0, 13.0, 24.0, 10.0

D_{low}: 20.0; 8.0, 22.5, 7.0, 4.0, 15.5, 22.5, 21.0, 12.0, 15.5, 19.0,

(2) For the low dose, the rank sum is 169, (3) the test statistic is:

$$ds_{D_{\text{low}}} = \frac{169 - 12 \cdot 25/2}{\sqrt{12 \cdot 12 \cdot 25/12}} = 0.87$$

and

$$ds_{D_{low}} = \frac{146 - 10 \cdot 23/2}{\sqrt{12 \cdot 10 \cdot 23/12}} = 2.04,$$

(4) with the quantile $d_{2,\infty,0.95} = 1.92$, the null hypothesis can be rejected for the comparison with the high dose only.

Assuming a priori monotonic ordering, a simple procedure on ordered t tests (each at level α) should be demonstrated (so-called pairwise contrasts):

$$t_j = \frac{\bar{x}_j - \bar{x}_C}{\sqrt{\text{MSE}(1/n_j + 1/n_C)}}$$

$$t_{D_{low}} = 0.756; \quad t_{D_{high}} = 2.744$$

The procedure starts with $t_{D_{high}}$. If $t_{D_{high}} > t_{\nu,1-\alpha'}$ proceed with the next lower dose; otherwise the procedure stops. Because $t_{D_{high}} = 2.744 > 1.696$, the global null hypothesis can be rejected and we proceed with D_{low}. Because $t_{D_{low}} = 0.756 < 1.696$, the procedure stops. This procedure guarantees experimental type I error and can be extended to any kind of pairwise contrast, that is, to nonparametric or dichotomous ones.

Moreover, we describe a simple score test for Lehmann alternatives for the groups C and D_{high} briefly. First the raw data are ranked into R_{ji} and transformed into

$$y_{ji} = (R_{ji}/(N + 1))^4,$$

where $N = n_1 + n_2$ denotes the global sample size. For example, the first value of D_{high} group is

$$R_{1,D_{high}} = 16 \longrightarrow y_{1,D_{high}} = (16/23)^4 = 0.234.$$

Using this transformed value, simple t tests are performed. In the example $\bar{y}_{D_{high}} = 0.330$, $s_{D_{high}} = 0.307$, and $\bar{y}_C = 0.068$, $s_c = 0.093$, and a p value of 0.0082 results. In this example, no mixing distribution of "responder and nonresponder" seems to exist because the p value of the t test on the raw data results in the same magnitude ($p = 0.0063$).

Analysis of Organ Weights

The specific problem in the analysis of organ weights is that changes in organ weights can be effected by changes in body weight (e.g., a decrease in liver organ weight can be caused by the test substance or by a decrease in body weight). The pragmatic solution for this problem is the analysis of both absolute and relative (organ weight/body weight) organ weights. From a statistical point of view, the analysis of this quotient assumes a linear relationship between the organ-versus-body weight and variance proportionality. Frequently, however, this is not the case (e.g., for the allometric function of organ weight = $a \cdot$ body weight^b, different values were observed for brain $b = 0.16$ and for prostate $b = 2.13$). Several alternative approaches have been published, including the analysis of covariance (with the covariate body weight). However, the simple concept of relative organ weights mentioned is routinely used.

Analysis of the Recovery Period

In some chronic studies, the test of reversibility becomes an important objective. Therefore, a subpopulation of at least one high-dose group and the control group are observed for some period of time (e.g., the recovery period) after the end of substance administration. For noninvasive measures (e.g., body weight and hematology parameters) paired tests should be used for the comparison between dosing and recovery periods in the same animals.

Analysis of Histopathological Findings

A biostatistical analysis of macroscopic and microscopic findings is frequently not done. However, selected histopathological findings may represent the primary endpoints. Moreover, a basic contradiction can be observed; the predictive value of easy and precise measurable variables is low, but it is high for the findings with the lowest possible information content: yes or no. Most of the findings are irreversible, but some clinical findings are reversible. For the latter, either the time of the first observation or the time period of the occurrence can be analyzed. The majority of the other findings can be analyzed as simple proportions. Here, the methodology of mortality rates (see the preceding section) or crude tumor rates (see the subsection Carcinogenicity Studies) can be used. Sometimes, histopathological findings are graded. Here, special tests for ordered categorical data should be used, preferably permutative versions.

Use of Reference Values

An inherent problem in the application of statistical significance tests in toxicology is the sometimes occurring contradiction between statistical significance and biological relevance. Several controversial approaches are discussed in the literature to overcome this problem. A simple approach is the additional use of reference or normal values. The rule is simple: first characterize what is "normal" and then classify the individual values into normal and nonnormal categories. A large proportion of nonnormal values will support the statistical significance.

TABLE 9 Number of Resorptions and Implantations in Control and D_{high}

		Litter									
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
C	r_{Ci}	2	2	1	0	0	1	1	0	1	0
	n_{Ci}	12	14	11	12	13	9	10	12	13	14
D	r_{Di}	5	3	3	2	7	3	1	4		
	n_{Di}	6	13	12	13	11	10	15	11		
Transformation to $y_{ji} = r_{ji}/n_{ji}$											
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
C	y_{Ci}	0.167	0.143	0.091	0	0	0.111	0.100	0	0.07	0
D	y_{Di}	0.833	0.231	0.250	0.154	0.636	0.300	0.067	0.364	7	

In regulatory toxicology, relatively good conditions exist for the estimation of normal values, since a large number of negative control animal data are available under standardized conditions (e.g., animals' housing, and analysis devices). From a statistical point of view, reference values can be categorized into those for a single future value, for a group (e.g., by the median) and for s values out of n_j of a group.

Here, only nonparametric reference values will be considered since the well-known parametric ones $[\bar{x} - 3s; \bar{x} + 3s]$ are very sensitive to a violation of the normal distribution assumption. A two-sided nonparametric reference region, including with $(1 - \alpha)$ probability $p100\%$ of the single values X_{ji} of an actual experiment, is

$$[X_{(f)}, X_{(n-f+1)}].$$

The parameter f can be found in Table 8 and the $X_{(\,)}$ represents the ordered single values of the historical controls from the smallest to the largest. The reference region is simply the f-smallest and the $(n - f + 1)$-largest value of this sample. The decision nonnormal is simply if $X_{ji} < X_{(f)}$ or $X_{ji} > X_{(n-f+1)}$ occurs.

REPRODUCTIVE TOXICOLOGY

For simplicity, the scheme of the most widely used teratogenicity study for estimation of blasto- and embryotoxic and teratogenic effects is presented in Fig. 3. From this scheme, some particularities found in comparison to the other toxicological studies follow.

- A hierarchical design: female ⊃ litter ⊃ fetus finding.
- The experimental unit "female" is not completely randomized according to the procedure of mating.
- The application of the substance occurs to the randomized sample unit "female," but most of the findings are on the fetuses.
- Numerous interactions occur (e.g., between the fetuses within a litter or concurrence between fetal mortality and the malformation rate, that is, the malformation rate could be biased by a large mortality rate).
- The primary endpoints are only discrete (e.g., the number of malformed fetuses per global number of fetuses).

Three statistical approaches can be distinguished according to these particularities:

- *Per fetus analysis:* Assume the fetus is the experimental unit. The independence assumption is violated and the false-positive rate is increased dramatically because of the pseudo increase in sample size. This approach should not be recommended.
- *Per litter analysis:* Assume the litter as the experimental unit, taking into account the interactions between the fetuses. This approach fits the biological background; however, the statistical methods (e.g., beta-binomial model) are complicated and frequently not robust.
- *Quasi per litter analysis:* This simple compromise is based on the transformation of the findings per litter $y_{ji} = r_{ji}/n_{ji}$, followed by a nonparametric test. Table 9 demonstrates this approach for the number of resorptions r_{ji} in relation to the number of implantations per litter (for control and high dose only). According to the transformed values y_{ij}, the permutative u test was used and a one-sided p value = 0.001 was found; that is, the rate of resorptions in the high dose is significantly increased with respect to the control.

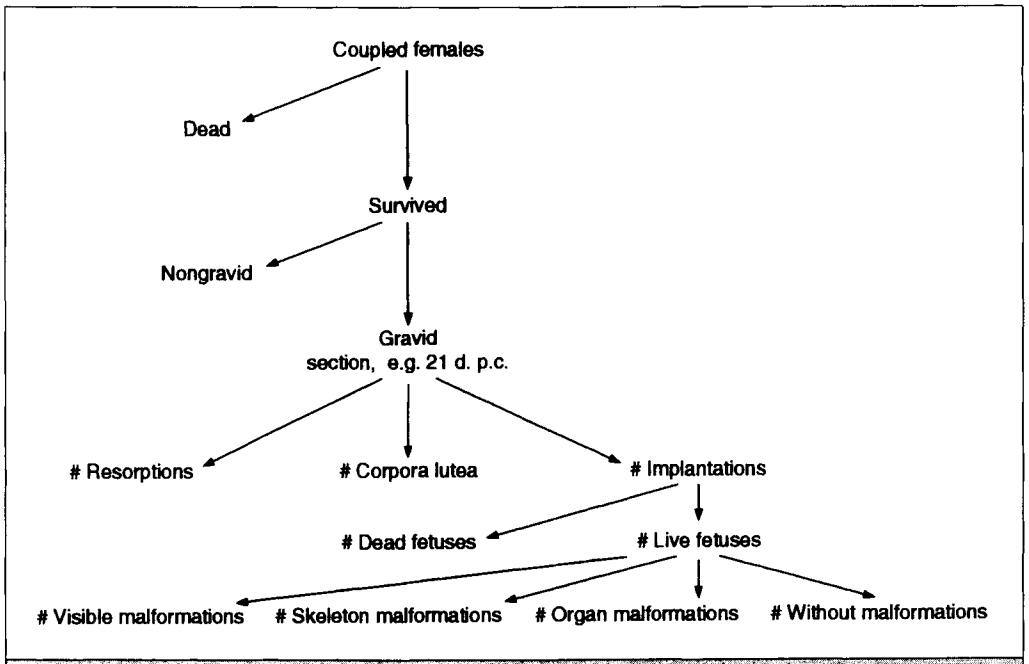


FIGURE 3 Scheme of teratogenicity studies (d.p.c. indicates days postcoito).

MUTAGENICITY STUDIES

To determine a potential mutagenic effect of a test substance, a testing battery consisting of several assays (e.g., salmonella–microsome (Ames), micronucleus, chromosome aberration, or HGPRT) is used. The most commonly used standardized assay is the salmonella–microsome (Ames) assay; hence its statistical analysis will be discussed as an example.

The objective of this assay is the determination of a potential gene-mutagenic effect of a substance after the single administration of several doses on histidine-free agar plates with specially modified *Salmonella typhimurium* bacteria strains. The numbers of grown visible colonies (revertants) of the dose groups are compared with those of the negative and positive controls. Several strains are used with different sensitivities for the different DNA mutagenic mechanisms, for example, base change mutation (TA 100 strain) or frameshift mutation (TA 98 strain).

TABLE 10 Biostatistical Analysis of Salmonella–Microsome (Ames) Assay Data on TA 100

j	Dose	Revertants	$\bar{x}_j$
0	0	80, 83, 79	80.7
1	0.01	92, 88, 78	86.0
2	0.02	96, 102, 98	98.7
3	0.04	104, 103, 95	100.7
4	0.08	94, 97, 99	96.7
5	0.16	93, 87, 91	90.3

The design is a completely randomized one-way layout $[C-, D_1, \ldots, D_k, C+]$, where the number of doses is between 3 and 6. Because the false-negative rate increases dramatically with too-low doses, the doses must be selected so that a nonmonotonic dose–response relationship occurs due to the interaction of mutagenic and cell-toxic effects. Therefore, the shape of the dose–response must be assumed to be an "umbrella."

Three statistical approaches can be distinguished:

- A formal twofold rule. That is, the effect is relevant if the mean values of two consecutive doses or the highest nontoxic dose are greater than 2 × the mean of the negative control (see, e.g., Cariello and Piegorsch 1996).
- A dose–response model. An example is the "single hit model,"

$$p(D_j) = (1 - \exp(-(a_1 + a_2 D_j)))\exp(-a_3 D_j),$$

  where the parameters a_1, a_2, and a_3 are to be estimated from the data.
- A suitable trend test that is robust against nonmonotonicity at high doses as well as the distribution of the revertants.

TABLE 11 Selected 95% Quantiles of the Jonckheere Trend Test[a]

Number of groups (incl. C)	n = const = 3	n = const = 4	n = const = 5
3	22	36	54
4	40	67	100
5	62	104	160
6	90	154	234

[a]Lehmann (1975).

In several publications, the distribution of the number of revertants is used in parametric models. However, even for the historical controls, this is difficult. Therefore, we use nonparametric tests for the discrete (including tied) endpoints with small sample sizes, called permutative versions. Simpson and Margolin (1986) published a Jonckheere trend test with a predetermined change point, while a simple partial order procedure using u tests can be found in Neuhäuser and Hothorn (1995).

Both the twofold and Simpson and Margolin (1986) approaches with Salmonella–microsome (Ames) assay data are demonstrated in Table 10. For the twofold rule there is no effect because there is no mean $\bar{x}_j > 161.4$ ($= 2 \cdot 80.7$). For the Simpson and Margolin (1986) approach, the change point q is calculated as:

TABLE 12 Mortality Data from a Carcinogenicity Study on Rats with a Duration of Two Years[a]

Control, $n_0 = 50$			D_{high}, $n_0 = 50$		
Week	Died	$W_C(t)$	Week	Died	$W_{D_{high}}(t)$
0	—	1.0	0	—	1.0
52	1	0.98	44	3	0.94
53	1	0.96	53	4	0.86
63	2	0.92	56	1	0.84
75	1	0.90	62	1	0.82
94	2	0.86	63	2	0.78
99	3	0.80	75	3	0.72
101	1	0.78	76	1	0.70
102	4	0.70	83	2	0.66
103	2	0.66	91	2	0.62
			94	4	0.54
			95	1	0.52
			99	3	0.48
			101	1	0.46
			102	1	0.44
			103	4	0.36

[a]Control and high dose only.

$$q = \max_{h \le j \le k} k \ge 1: \sum_{h=C}^{j} U_{hj} > c_j \quad \text{with } c_j = 0.5 n_j N_{j-1}$$

where U_{hj} denotes the Mann–Whitney counts for group h vs group j and $N_j = n_j$.

j	U	$\sum U_{hj}$	c_j	*Result*
1	U_{C1}	3	4.5	—
2	$U_{C2} + U_{12}$	18	9	
3	$U_{C3} + U_{13} + U_{23}$	24	13.5	
4	$U_{C4} + U_{14} + U_{24} + U_{34}$	23	18	max j
5	$U_{C5} + U_{15} + U_{25} + U_{35} + U_{45}$	15	22.5	

Here, the change point is $q = 4$. The Jonckheere trend test until q is

$$JT^q = \sum_{i=C}^{q-1} \sum_{j=i+1}^{q} U_{ij}$$

$$\begin{aligned} JT^q &= U_{C1} + U_{C2} + U_{C3} + U_{C4} + U_{12} + U_{13} + U_{14} \\ &\quad + U_{23} + U_{34} \\ &= 3 + 9 + 9 + 9 + 9 + 9 + 9 + 6 \\ &\quad + 3 + 2 = 68. \end{aligned}$$

This test value must be compared with the quantile $r_{k=5,n=3,1-\alpha=0.95} = 61$. Selected values for the quantiles of the Jonckheere trend test are given in Table 11. Because $JT^q = 68 > 62$, the null hypothesis is rejected (p value of the permutative test $p = 0.004$). If the Jonckheere trend test had been used for all doses, ignoring the nonmonotonicity,

$$JT = \sum_{i=C}^{k-1} \sum_{j=i+1}^{k} U_{ij}$$

$$\begin{aligned} JT &= U_{C1} + U_{C2} + U_{C3} + U_{C4} + U_{C5} + U_{12} \\ &\quad + U_{13} + U_{14} + U_{15} + U_{23} + U_{24} + U_{25} \\ &\quad + U_{34} + U_{35} + U_{45} \\ &= 3 + 9 + 9 + 9 + 9 + 9 + 9 + 9 + 6 \\ &\quad + 6 + 3 + 0 + 2 + 0 + 0 = 83 \end{aligned}$$

the null hypothesis would not be rejected (quantile $r_{k=6,n=3,1-\alpha=0.95} = 90$; see Table 11).

TABLE 13 Analysis of Tumors without Mortality Adjustment: Males with Benign Leydig-Cell Tumors

	C	1 mg/kg	3 mg/kg	10 mg/kg
$n_{\text{Leydig cell tumor}}$	1	4	3	5
n_0	100	50	50	50
$n_{\text{autolysis}}$	1	0	0	0
$n_{\text{died before 6 months}}$	2	1	0	2
$n_{\text{under tumor risk}}$	97	49	50	48

CARCINOGENICITY STUDIES

The objective of a carcinogenicity study is the estimation of a carcinogenic effect over a lifetime administration of a substance. The long life is accomplished by the use of specific pathogen-free animals, germ-free environment, standard diet, constant temperature, and constant moisture in animals' housing. In this discussion, only terminal killing will be assumed. Whereas serial sacrifies allow for more precise conclusions, the statistical methodology is complicated and beyond the scope of this discussion.

The carcinogenic effect can be determined by:

- Comparing the tumor incidences (or tumor–time relationships) between dose and control groups (primary objective).
- Comparing the survival functions between dose and control groups (secondary objective).
- Comparing additional measures, like body weight, food consumption, hematology, clinical chemistry, and nonneoplastic histopathological findings (tertiary objective).

We will only discuss the primary and secondary objectives. The analysis of the additional measures can be done in analogy to the chronic studies.

Analysis of Mortality

On the one hand, dose-dependent mortality is a criterion for substance effect; on the other hand, mortality and tumor growth are correlated; that is, mortality is the inherent effect of extensive tumor growth, but non-tumor-caused mortality is a competitor to tumor formation.

First for all animals, the cause of death must be evaluated (sometimes by histopathological examination):

1. Died according to the study protocol
2. Killed in extremes (moribund)
3. Killed according to the study protocol
4. Died not according to the study protocol (e.g., artifacts or erroneous manipulation during narcosis or gavaging)

Animals of category 4 will be removed before the analysis; animals of categories 2 and 3 are censored (i.e., nominal date of death ≠ unknown spontaneous date of death). The statistical problem consists in the comparison of the survival functions between dose groups and the control. Frequently, a dose trend can be assumed. Model-based (e.g., proportional-hazard model) or model-free (e.g., nonparametric tests for censored data using generalized Wilcoxon statistics) assumptions can be used.

Graphical representation of the product–limit estimator of the survival functions (and if possible the confidence curves) is helpful (Table 12). Statistical differences between these survival-time functions can be estimated by confidence intervals or special tests for censored data.

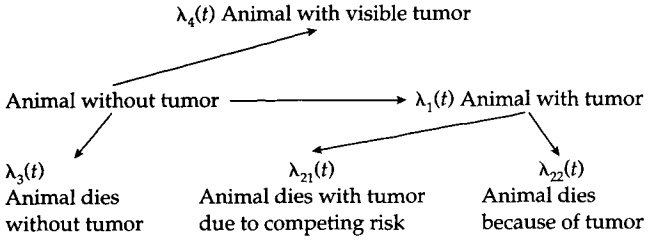
Kaplan and Meier (1958) Estimator

$$W_j(t) = \prod_k (1 - x_{jk}/n_{jk})$$

where t denotes the different times t_k (e.g., 52, 53, or 63 weeks), x_{jk} is the number of death in group j at time t_k, and n_{jk} is the number of animals in group j at risk of dying at t_k. For example, for the high-dose-group results until week 44, $W_{D_{\text{high}}}(1) = (1 - 3/50) = 0.94$; until week 53, $W_{D_{\text{high}}}(2) = 0.94*(1 - 4/47) = 0.86$.

Analysis of Tumors

The principles of tumor development can be illustrated by



where $\lambda_i(t)$ denotes the probabilities from one level to another. The definition of "carcinogenic risk" is a multisource problem.

- Increased incidence of a dose group with respect to the negative control (increased tumor risk $\lambda_1(t)$).

- Increased prevalence of a dose group with respect to the negative control, where prevalence is the probability (tumor exists | animal is living).
- Reduction of latency time of a tumor of a dose group with respect to the negative control where latency time is $t \mid \lambda_1(t) = 0$.
- Increased incidence of visible tumors (i.e., increased risk of $\lambda_4(t)$).
- Observation of untypical tumor sites for the used strain.
- Increased degree of metastazation or multiple tumors within one site.

Tumors should be classified as:

- Mortality-independent tumors (e.g., tumors that are visible or found palpation). The latency time can be predicted with more or less precision.
- Internal tumors found during macroscopic or microscopic examination.

1. fatal tumors (i.e., cause of the animal's death)
2. incidental tumors (i.e., not the cause of the animal's death)

Competing risks could be other tumors, non-neoplastic reasons, or preprotocol sacrifice. Other classification criteria are

- Benign tumors.
- Malignant tumors.
- Tumors of an organ system (e.g., tumors in the sexual system).
- Tumors of an organ (e.g., liver tumors).
- Tissue-specific organs (e.g., bronchial adenoma).

These tumor findings can be analyzed using mortality-unadjusted (crude proportions) or mortality-adjusted methods. Independent of the magnitude of the global mortality differences, a mortality-adjusted analysis is necessary. Crude proportions (percentages) can be used; however, the statistical results might be biased.

TABLE 14 Analysis of an Incidental Tumor

Time interval	Definition	C	D_{low}	D_{high}
81–104 weeks	Number of tumors	0	0	2
	Number of animals at tumor risk	15	14	19
Terminal sacrifice	Number of tumors	0	1	0
	Number of animals at tumor risk	37	31	26

Analysis of Tumor Incidences without Mortality Adjustment

Here, the data of a $2 \times (k + 1)$ contingency table are analyzed.

	Control	*Dose*$_1$	⋯	*Dose*$_k$	*Total*
Number of animals with tumor	r_C	r_{D1}	⋯	r_{Dk}	$r.$
Number of animals without a tumor	o_C	o_{D1}	⋯	o_{Dk}	$o.$
$\sum$ = animals under tumor risk	n_C	n_{D1}	⋯	n_{Dk}	$n.$

Usually, separate pairwise comparisons of treated animals versus the negative control are performed by a one-sided approximate χ^2 two-sample test for proportions with continuity correction. Because of the small sample sizes and tumor rates, an exact test should be done (e.g., Fisher's or Barnard's test). Even more appropriate are many-to-one procedures for proportions; see the Passing (1984) procedure.

The correct definition of the number of animals under tumor risk is important. Normally, from the initial sample sizes n_0, the number of animals that died very early (i.e., before a predefined time, such as 12 months or before the first tumor was observed will be subtracted). This will also be done for the animals that are not autopsied (because, e.g., of autolysis or cannibalism). Therefore, the organ-specific number under risk is normally different.

Because the decision for a biologically relevant tumor effect is valid if a monotonic dose–response exists, a trend test should be used. Here again, the Armitage (1955) trend test can be used. Tumor incidences are analyzed in this way (Table 13). Because of small animal and tumor numbers, the exact version of the Armitage (1955) trend test was used.

The one-sided p values is 0.016; hence a dose-dependent increase of Leydig cell tumor can be assumed.

In a carcinogenicity study, multiple tumor sites are to be analyzed simultaneously. Up to about 30 to 50 sites are examined because an a priori definition of a target tumor site is seldom possible. Therefore, a multiple-testing problem exists. Some statistical procedures exist to treat this problem, but a simple decision rule according to Haseman (1983) can also be used: A relevant effect is valid if the p value is less than 0.05

for a common tumor (spontaneous rate > 1%); the p value should be less than 0.025 for a rare tumor (spontaneous rate < 1%).

Tumor findings can also be reported as total number of tumors, number of benign tumors, number of malignant tumors, number of tumors of an organ system, and the number of tumors in an organ. Again, the increased false-positive rate should be taken into account.

Mortality-Adjusted Tumor Analysis

For fatal tumors, the null hypothesis assumes that the risk of dying is equal to the risk of developing a tumor and that this risk is the same for all groups. Therefore, all surviving animals and animals that died because of a competing risk are censored. Data should be simply arranged in a stratified $2 \times (k + 1)$ contingency table, where the strata represent the order of mortality time (e.g., as experimental weeks). The responder r_{ji} is the number of animals with a selected fatal tumor (i.e., tumor-site specific or pooled) and n_{ji} is the number of animals at tumor risk in this time interval.

The risks of mortality for animals without a tumor and those with an incidental tumor are assumed to be equal. The null hypothesis is: All animals within a selected time interval have the same probability of an incidental tumor regardless of the treatment. Therefore, an analogous stratified table similar to the fatal-tumor table exists. However, the number of animals at tumor risk in a particular time period is the number of animals that died or were sacrificed in that time period. Suitable time intervals should be chosen so that the number of animals at tumor risk is not zero for any treatment group. This can be accomplished by methods that are data-driven (ad hoc runs according to Peto *et al.* 1980) or with an a priori definition (e.g., for rats, time intervals should be defined according to the National Toxicology Program (NTP) into 0–12, 12–18, 18–20, 20–22, and 22–24 months and terminal sacrifice intervals). The latter is much simpler and more reproducible, and is recommended. Combined analysis of both fatal and incidental tumors is also possible. In Table 14, a real data example is analyzed in this way. The animals at tumor risk in the time interval are those animals that died or were killed. Because of the small number of animals at risk and the small number of tumors, a permutative test is recommended. Here a stratified Armitage (1955) trend test was used and a p value of 0.07 (asymptotic version) was calculated by the program STATXACT (Metha *et al.*, 1995). Combined analysis of both fatal and incidental tumors is also possible.

BIOSTATISTICS IN REGULATORY TOXICOLOGY

The good laboratory practice (GLP) regulations have had a deep influence on the conduct of regulatory toxicology studies. The major principles, standard operating procedures (SOPs) and validation, are also applicable to the statistics. Hence, for biostatistical analysis, SOPs should be worked out and only validated software should be used.

Bibliography

Armitage, P. (1955). Tests for linear trends in proportions and frequencies. *Biometrics* **11,** 375–386.

Cariello, N. F., and Piegorsch, W. W. (1996). The Ames assay: The two-fold rule revisited. *Mutation Res.* **369,** 23–31.

Conover, W. J., and Salsburg, D. S. (1988). Locally most powerful tests for detecting treatment effects when only a subset of patients can be expected to respond to treatment. *Biometrics* **44,** 188–196.

Dunnett C. W. (1955). A multiple comparison procedure for comparing several treatments with a control. *J. Am. Statist. Ass.* **50,** 1096–1121.

Hahn G. J., and Meeker, W. Q. (1991). "Statistical Intervals." John Wiley and Sons, New York. pp. 75–99.

Haseman, J. (1983). A re-examination of false-positive rates for carcinogenicity studies. *Fund. Appl. Toxicol.* **3,** 334–339.

Hsu, J. C. (1996). "Multiple Comparisons." Chapman & Hall, New York.

Jonckheere, A. R. (1954). A distribution-free k-sample test against ordered alternatives. *Biometrika* **41,** 133–145.

Kaplan, E. L., and Meier, P. (1958). Nonparametric estimation from incomplete observations. *J. Am. Statist. Ass.* **53,** 457–481.

Lin, K. K., and Ali, M. W. (1994). Statistical review and evaluation of animal tumorigenicity studies. *In* "Statistics in the Pharmaceutical Industry" (C. R. Buncher and J. Y. Tsay, eds.), pp. 19–57. M. Dekker, New York.

Kruskal, W. H., and Wallis, W. A. (1952). Use rank on one-criterion variance analysis. *J. Am. Statist. Ass.* **47,** 583–621.

Lehmann, E. L. (1975). "Nonparametrics." Holden-Day, New York. p. 429.

Levene H. (1960). Robust tests for equality of variance. *In* "Contributions to Probability and Statistics," pp. 278–292. Stanford University Press.

Marcus, R., Peritz, E., and Gabriel, K. B. (1976). On closed testing procedures with special reference to ordered analysis of variance. *Biometrika.* **63,** 655–660.

Mehta C, and Patel N. (1995). StatXact 3 for Windows. Cytel Software Corp., Cambridge.

Neuhäuser, M., and Hothorn, L. A. (1995). Auswertung der Dosis-Wirkungs-Abhängigkeit des Ames Mutagenitätsassay bei direkter Kontrolle des Konsumentenrisikos. *In* "Medizin. Informatik, Biometrie und Epidemiologie, Vol. 80, Medizinische Forschung und ärztliches Handeln" MVM Medizin Verlag, München (H. J. Trampisch and S. Lange, eds.), pp. 113–116.

Neuhäuser, M, and Hothorn, L. A. (1997). Trend tests for dichotomous endpoints. *Drug Information J.* **30,** 463–469.

Passing, H. (1984). Exact simultaneous comparisons with control in a rxc contingency table. *Biometr. J.* **26,** 643–654.

Peto, R. *et al.* (1980). Guidelines for simple, sensitive significance tests for carcinogenic effects in long-term animal experiments. *In* "Long-Term and Short-Term Screening Assays for Carcinogens: A Critical Appraisal," IARC monographs supplement 2, pp. 311–426. IARC Lyon.

Simpson, D. G., and Margolin, B. H. (1986). Recursive nonparametric testing for dose-response relationship subject to downturns at high doses. *Biometrika* **73,** 589–596.

Steel, R. G. D. (1959). A multiple comparison sign test: Treatments vs. control. *J. Am. Statist. Ass.* **54,** 767–775.

This Page Intentionally Left Blank

CHAPTER

3

Principles of Toxicokinetics

BURCKHARD FICHTL

Walther Straub Institute of Pharmacology and Toxicology, Ludwig Maximilians University, Munich, Germany

INTRODUCTORY REMARKS

The effects of any foreign compound (xenobiotic) acting on the body of man or animals depend on the pharmacodynamic and pharmacokinetic properties of the substance. Pharmacodynamics (Greek *φαρμακον* = drug, poison; *δνναμισ* = strength, power) deals with the effects that a substance exerts and the mechanisms of action, and thus describes "what a substance does to the body." Both the magnitude and the time course of effects also depend on the fate of a substance in the body. Pharmacokinetics (Greek *κινειν* = to move) deals with the time course of the concentrations of a substance in the body, which is the net result of the interplay between absorption, distribution, and elimination. Pharmacokinetics thus describes "what the body does to the substance."

What Is Toxicokinetics?

According to Case (1993), the term "toxicokinetics" was apparently first used in a paper published by a Russian author in 1937. Since then, the term toxicokinetics has gained in popularity, but its use has not been very consequent (see the examples given in Table 1). In recent years, there appears to be a tendency to restrict its use to preclinical toxicology studies. It may be questioned, however, whether the peculiarities of these studies, such as the use of different animal species and rather high doses, merit toxicokinetics being regarded as a special discipline. For instance, kinetic aspects may be of importance also in toxicological studies concerned with the effect of low doses of environmental chemicals in humans. One should not lose sight of the fact that toxicokinetics—in any form—is essentially based on pharmacokinetic principles that apply to any xenobiotic, be it a drug used for therapeutical purposes or a toxicant. In order to avoid possible embarrassment, in the present chapter we will largely rely on the neutral term "kinetics."

Pharmco(toxico)kinetic Parameters

The fate of a drug (or any xenobiotic) in the body may be characterized by the LADME scheme (Fig. 1.) If a drug is administered in a drug product, it must be liberated first. If not injected directly into the circulation, a drug must be absorbed from an absorption site (e.g., gastrointestinal fluid or intramuscular or subcutaneous injection site). After entering the systemic circulation, the drug is distributed via the blood stream to the tissues and will eventually reach its site(s) of action. To a certain extent, most substances become reversibly bound to plasma proteins and/or tissue constituents. A substance may be degraded by metabolism (biotransformation) and ultimately it is excreted from the body. During the processes of absorption, distribution, metabolism, and excretion, a drug has to permeate biological membranes. Drugs cross membranes mainly by passive diffusion. According to Fick's law, the diffusion rate is directly proportional to the concentration gradient across a membrane. Furthermore, it depends on the lipid solubility of the substance in question. For an ionizable drug, lipid solubility depends on its pK_a value and the pH of body fluids. Obviously, the fate of a sub-