The Process of New Drug Discovery and Development

Second Edition

The Process of New Drug Discovery and Development

Second Edition

Edited by
Charles G. Smith Ph.D.
James T. O'Donnell Pharm.D.

informa
healthcare
New York London

Informa Healthcare USA, Inc.
270 Madison Avenue
New York, NY 10016

Printed in the United States of America on acid-free paper
10 9 8 7 6 5 4 3 2 1

International Standard Book Number-10: 0-8493-2779-2 (Hardcover)
International Standard Book Number-13: 978-0-8493-2779-7 (Hardcover)
Library of Congress Card Number 2005036791

Library of Congress Cataloging-in-Publication Data

The process of new drug discovery and development / editors, Charles G. Smith and James T. O'Donnell.-- 2nd ed.
p. ; cm.
Rev. ed. of: The process of new drug discovery and development / Charles G. Smith. c1992.
Includes bibliographical references and index.
ISBN-13: 978-0-8493-2779-7 (alk. paper)
ISBN-10: 0-8493-2779-2 (alk. paper)
1. Drugs--Research--History. 2. Drugs--Design--History.
[DNLM: 1. Drug Design. 2. Drug Evaluation, Preclinical. 3. Drug Evaluation. QV 744 P9645 2006] I. Smith, Charles G. (Charles Giles) II. O'Donnell, James T., Pharm. D.

RM301.25.S55 2006
615'.19--dc22 2005036791

Visit the Informa Web site at
www.informa.com

and the Informa Healthcare Web site at
www.informahealthcare.com

As was the case with the first edition, this book is dedicated to my irreplaceable wife, Angeline, who persisted in encouraging me in this and other scientific efforts that often left little time for recreational activities. In addition, I dedicate it to my daughter, Tracy, whose strong encouragement and occasional harassment convinced me to write the first edition.

Charles G. Smith

To my wife, Sylvia, and my children, Kimberly and Jim, who make my life worth living.

James T. O'Donnell

Acknowledgments

Sincere appreciation is expressed to Mr. S. Zollo at CRC Press/Taylor & Francis, whose strong solicitation resulted in my decision to participate in the second edition of this book. The role of my coeditor, Dr. James T. O'Donnell, without whose expertise and organizational abilities this work could not possibly have been accomplished, deserves my most heartfelt thanks and appreciation. The participation of many excellent authors has made this book, in my opinion, a truly valuable contribution to the scientific literature in the field of drug discovery and development. I am honored to end a most enjoyable career with this contribution.

Charles G. Smith

Editors

Charles G. Smith earned the B.S. degree (1950) in chemistry from the Illinois Institute of Technology, the M.A. degree (1952) in biochemistry from Purdue University, and the Ph.D. degree (1954) in biochemistry from the University of Wisconsin. He joined The Upjohn Company in 1954 and worked in fermentation biochemistry for several years. In 1962, he was appointed head of the biochemistry department at Upjohn and initiated a major anticancer effort therein. Dr. Smith moved to E. R. Squibb & Sons pharmaceutical company in 1968, where he became vice president for research. He joined the Revlon Health Care Group in 1975 as vice president for research and development. In 1986, he retired from industry and became a pharmaceutical consultant. During his tenure in the major pharmaceutical companies, Dr. Smith was intimately involved with projects in the fields of infectious diseases, cancer, cardiovascular diseases, central nervous system diseases, and pharmaceutical products from blood plasma. Since 1986, he has consulted with many biotechnology companies that work in a broad cross-section of pharmaceutical research. He was a cofounder of Vanguard Medica in the United Kingdom in 1991 and named adjunct professor in the Department of Natural Sciences at San Diego State University in the same year. Dr. Smith is the author of 49 publications and the first edition of this book (1992), and remains a pharmaceutical consultant in the biotechnology field.

James T. O'Donnell earned the B.S. degree in pharmacy from the University of Illinois (1969) and the Doctor of Pharmacy degree from the University of Michigan (1971) as well as the M.S. degree in clinical nutrition from Rush University (1982). He completed a residency in clinical pharmacy at the University of Illinois Research Hospitals and has been a registered pharmacist in Illinois since 1969. Dr. O'Donnell spent 17 years in clinical practice at both the Cook County Hospital and the Rush University Medical Center in Chicago. Dr. O'Donnell is currently an associate professor of pharmacology at the Rush University Medical Center and is a member of the Graduate College, involved in the teaching of new drug development and regulations. Also, Dr. O'Donnell is a lecturer in the Department of Medicine at the University of Illinois College of Medicine, Rockford. He is a Diplomate of the American Board of Clinical Pharmacology and Board of Nutritional Specialties, a fellow of the American College of Clinical Pharmacology and the American College of Nutrition, and a member of several professional societies. Dr. O'Donnell is the author of 257 publications, and he is the founding editor of the *Journal of Pharmacy Practice*, a co-editor of *Pharmacy Law*, and the editor of *Drug Injury: Liability, Analysis, and Prevention, First and Second Editions*. In addition to his academic and editorial endeavors, Dr. O'Donnell regularly consults in drug and pharmaceutical matters to industry, government, and law, and serves as pharmacologist consultant to the State of Illinois Department of Public Health.

Contributors

Loyd V. Allen University of Oklahoma College of Pharmacy, Edmond, Oklahoma, U.S.

Donald C. Anderson Pharmacia Corporation, Kalamazoo, Michigan, U.S.

Timothy Anderson Pfizer, Inc., Groton, Connecticut, U.S.

Stephen Barrett Quackwatch, Inc., Allentown, Pennsylvania, U.S.

Joanne Bowes Safety Assessment UK, AstraZeneca R&D Alderley Park, Macclesfield, Cheshire, U.K.

Irwin A. Braude Compass Pharmaceuticals, LLC, Charlotte, North Carolina, U.S.

Celia Brazell Genetics Research, GlaxoSmithKline, Greenford, Middlesex, U.K.

Jerry Collins Food and Drug Administration, Center for Drug Evaluation and Research, Office of Testing and Research, Rockville, Maryland, U.S.

William T. Comer TorreyPines Therapeutics, Inc., La Jolla, California, U.S.

Mark Crawford Cerep, Redmond, Washington, U.S.

Andrew Dorner Wyeth Research, Andover, Maryland, U.S.

David Essayan Food and Drug Administration, Center for Biologics Evaluation and Research, Office of Therapeutics, Division of Clinical Trial Design and Analysis, Rockville, Maryland, U.S.

Pauline Gee Department of Predictive Biology, MDS Pharma Services, Bothell, Washington, U.S.

Baltazar Gomez-Mancilla Pharmacia Corporation, Kalamazoo, Michigan, U.S.

Henry Grabowski Duke University, Durham, North Carolina, U.S.

Howard E. Greene Rancho Santa Fe, CA

Joseph Hackett Food and Drug Administration, Center for Devices and Radiological Health, Office of In Vitro Diagnostic Device Evaluation and Safety, Rockville, Maryland, U.S.

Jeff M. Hall Department of Cell Biology, Genoptix, Inc., San Diego, California, U.S.

Tim G. Hammond Safety Assessment UK, AstraZeneca R&D Alderley Park, Macclesfield, Cheshire, U.K.

Valérie Hamon Cerep, Paris, France

Ismael J. Hidalgo Absorption Systems, LP, Exton, Pennsylvania, U.S.

Carol Ann Homon Boehringer Ingelheim Pharmaceuticals, Inc., Medicinal Chemistry Department, Ridgefield, Connecticut, U.S.

Shiew-Mei Huang Food and Drug Administration, Center for Drug Evaluation and Research, Office of Clinical Pharmacology and Biopharmaceutics, Rockville, Maryland, U.S.

Susan Ide National Institute of Health, NHGRI, and Novartis, Gaithersburg, Maryland, U.S.

Vincent Idemyor College of Medicine, University of Illinois, Chicago, Illinois, U.S.

Thierry Jean Cerep, Paris, France

Ilona Kariv Applications Research and Development, Genoptix Inc., San Diego, California, U.S.

Joanne Killinger Wyeth Research, Chazy, New York, U.S.

Harold J. Kwalwasser MDS Pharma Services, Bothell, Washington, U.S.

John Leighton Food and Drug Administration, Center for Drug Evaluation and Research, Office of New Drugs, Rockville, Maryland, U.S.

Lawrence J. Lesko Food and Drug Administration, Center for Drug Evaluation and Research, Office of Clinical Pharmacology and Biopharmaceutics, Rockville, Maryland, U.S.

Jibin Li Absorption Systems, LP, Exton, Pennsylvania, U.S.

Elizabeth Mansfield Food and Drug Administration, Center for Devices and Radiological Health, Office of In Vitro Diagnostic Device Evaluation and Safety, Rockville, Maryland, U.S.

Phillip J. Marchand Optical Systems, Genoptix Inc., San Diego, California, U.S.

Patricia A. McNeeley Genoptix Inc., San Diego, California, U.S.

Robert Meyer Food and Drug Administration, Center for Drug Evaluation and Research, Office of New Drugs, Rockville, Maryland, U.S.

Justina A. Molzon U.S. Food and Drug Administration, Center for Drug Evaluation and Research, Rockville, Maryland, U.S.

Daniel Mufson Apotherx LLC, Napa, California, U.S.

Richard M. Nelson Boehringer Ingelheim Pharmaceuticals, Inc., Medicinal Chemistry Department, Ridgefield, Connecticut, U.S.

Lori Nesbitt Discovery Alliance International, Inc., Sugarland, Texas, U.S.

Sarfaraz K. Niazi Pharmaceutical Scientists, Inc., Deerfield, Illinois, U.S.

F. Richard Nichol Nichol Clinical Technologies Corporation, Newport Beach, California, U.S.

Tina S. Nova Genoptix Inc., San Diego, California, U.S.

James T. O'Donnell Department of Pharmacology, Rush Medical College, Rush University Medical Center, Chicago, IL

Mark Paich Lexidyne, LLC, Colorado Springs, Colorado, U.S.

Corey Peck Lexidyne, LLC, Colorado Springs, Colorado, U.S.

Madhu Pudipeddi Technical Research and Development, Novartis Pharmaceuticals Corporation, Mumbai, India

Michael H. Rabinowitz Johnson & Johnson Pharmaceutical Research & Development LLC, San Diego, California, U.S.

Mitchell E. Reff Biogen Idec, Inc., Oncology, Cambridge, Massachusetts, U.S.

Frederick E. Reno Reno Associates, Merritt Island, Florida, U.S.

Michael G. Rolf Safety Assessment UK, AstraZeneca R&D Alderley Park, Macclesfield, Cheshire, U.K.

Stephen G. Ryan AstraZeneca Pharmaceuticals, Wilmington, Delaware, U.S.

Ronald A. Salerno WorldWide Regulatory Affairs, Wyeth Research, St. Davids, Pennsylvania, U.S.

Richard H.C. San Genetic Toxicology, BioReliance, Rockville, Maryland, U.S.

Virginia Schmith GlaxoSmithKline, Research Triangle Park, North Carolina, U.S.

Abu T.M. Serajuddin Novartis Pharmaceutical Corporation, Pharmaceutical and Analytical Development, East Hanover, New Jersey, U.S.

Nigel Shankley Johnson & Johnson Pharmaceutical Research & Development, LLC, Merryfield Row, San Diego, California, U.S.

Peter Shaw Department of Pharmacogenomics and Human Genetics, Bristol-Meyers Squibb, Princeton, New Jersey, U.S.

Frank Sistare Food and Drug Administration, Center for Drug Evaluation and Research, Office of Testing and Research, Rockville, Maryland, U.S.

Charles G. Smith Rancho Santa Fe, CA

Kirk Solo Lexidyne, LLC, Colorado Springs, Colorado, U.S.

John C. Somberg Chief Division of Clinical Pharmacology, Department of Medicine & Pharmacology, Rush University, Chicago, Illinois, U.S.

Brian B. Spear Abbott Laboratories, Department of Pharmacogenetics, Abbott Park, Illinois, U.S.

Jason Valant Lexidyne, LLC, Colorado Springs, Colorado, U.S.

Jean-Pierre Valentin Safety Assessment UK, AstraZeneca R&D Alderley Park, Macclesfield, Cheshire, U.K.

Gönül Veliçelebi Research and Drug Discovery, TorreyPines Therapeutics**,** Inc., La Jolla, California, U.S.

Daniel D. Von Hoff Translational Genomics Institute, Phoenix, Arizona, U.S.

Heather L. Wallace Scientific Communications, ICON Clinical Research, North Wales, Pennsylvania

Mark Watson Clinical Genomics, Merck and Co., Inc., West Point, Pennsylvania, U.S.

Janet Woodcock U.S. Food and Drug Administration, Center of Drug Evaluation and Research, Rockville, Maryland, U.S.

Alexandra Worobec Food and Drug Administration, Center for Biologics Evaluation and Research, Office of Therapeutics, Division of Clinical Trial Design and Analysis, Rockville, Maryland, U.S.

Contents

1

Introduction

Charles G. Smith

Prior to the 20th century, the discovery of drug substances for the treatment of human diseases was primarily a matter of "hit or miss" use in humans, based on folklore and anecdotal reports. Many, if not most, of our earliest therapeutic remedies were derived from plants or plant extracts that had been administered to sick humans (e.g., quinine from the bark of the cinchona tree for the treatment of malaria in the mid-1600s and digitalis from the foxglove plant in the mid-1700s for the treatment of heart failure, to name two). Certainly, some of these early medications were truly effective (e.g., quinine and digitalis) in the sense that we speak of effective medications today. On the other hand, based on the results of careful studies of many such preparations over the years, either in animals or man, one is forced to come to the conclusion that most likely, the majority of these plant extracts was not pharmacologically active, but rather they were perceived as effective by the patient because of the so-called placebo effect. Surprisingly, placebos (substances that are known not to be therapeutically efficacious, but that are administered so that all the psychological aspects of consuming a "medication" are presented to the patient) have been shown to exert positive effects in a wide range of disease states, attesting to the "power of suggestion" under certain circumstances. There still exist today practitioners of so-called homeopathic medicine, which is based on the administration of extremely low doses of substances with known or presumed pharmacologic activities. For example, certain poisons, such as strychnine, have been used as a "tonic" for years in various countries at doses that are not only nontoxic but that in the eyes of most scientifically trained medical and pharmacological authorities, could not possibly exert an actual therapeutic effect. Homeopathy is practiced not only in underdeveloped countries, but also in certain well-developed countries, including the United States, albeit on a very small scale. Such practices will, most likely, continue since a certain number of patients who require medical treatment have lost faith, for one reason or another, in the so-called medical establishment. More will be said about proving drug efficacy in Chapters 8 to 10.

Pioneers in the field of medicinal chemistry such as Paul Ehrlich (who synthesized salvarsan, the first chemical treatment for syphilis, at the turn of the 20th century), were instrumental in initiating the transition from the study of plants or their extracts with purported therapeutic activities to the deliberate synthesis, in the laboratory, of a specific drug substance. Certainly, the discovery of the sulfa drugs in the 1930s added great momentum to this concept, since they provided one of the earliest examples of a class of pure chemical compounds that could be unequivocally shown to reproducibly bring certain infectious diseases under control when administered to patients by mouth. During World War II,

the development of penicillin stimulated an enormous and highly motivated industry aimed at the random testing (screening) of a variety of microbes obtained from soil samples for the production of antibiotics. This activity was set into motion by the discovery of Alexander Fleming and others in England in 1929 that a *Penicillium* mold produced tiny amounts of a substance that was able to kill various bacteria that were exposed to it in a test tube. When activity in experimental animal test systems and in human patients was demonstrated, using extremely small amounts of purified material from the mold broth (penicillin), it was immediately recognized that antibiotics offered a totally new route to therapeutic agents for the treatment of infectious diseases in human beings. In addition to the scientific interest in these findings, a major need existed during World War II for new medications to treat members of the armed forces. This need stimulated significant activity on the part of the United States Government and permitted collaborative efforts among pharmaceutical companies (which normally would be highly discouraged or prohibited by antitrust legislation from such in-depth cooperation) to pool resources so that the rate of discovery of new antibiotics would be increased. Indeed, these efforts resulted in accelerated rates of discovery and the enormous medical and commercial potential of the antibiotics, which were evident as early as 1950, assured growth and longevity to this important new industry. Major pharmaceutical companies such as Abbott Laboratories, Eli Lilly, E. R. Squibb & Sons, Pfizer Pharmaceuticals, and The Upjohn Company in the United States, to name a few, were particularly active in these endeavors and highly successful, both scientifically and commercially, as a result thereof (as were many companies in Europe and Japan). From this effort, a wide array of new antibiotics, many with totally unique and completely unpredictable chemical structures and mechanisms of action, became available and were proven to be effective in the treatment of a wide range of human infectious diseases.

In the 1960s and 1970s, chemists again came heavily into the infectious diseases' arena and began to modify the chemical structures produced by the microorganisms, giving rise to the so-called semi-synthetic antibiotics, which form a very significant part of the physicians' armamentarium in this field today. These efforts have proved highly valuable to patients requiring antibiotic therapy and to the industry alike. The truly impressive rate of discovery of the 'semi-synthetic' antibiotics was made possible by the finding that, particularly in the penicillin and cephalosporin classes of antibiotics, a portion of the entire molecule (the so-called 6-APA in the case of penicillin and 7-ACA in the case of cephalosporin) became available in large quantities from fermentation sources. These complex structures were not, in and of themselves, able to inhibit the growth of bacteria, but they provided to the chemist the central core of a very complicated molecule (via the fermentation process), which the chemist could then modify in a variety of ways to produce compounds that were fully active (hence the term 'semi'-synthetic antibiotics). Certain advantages were conferred upon the new molecules by virtue of the chemical modifications such as improved oral absorption, improved pharmacokinetic characteristics and expanded spectrum of organisms that were inhibited, to name a few. Chemical analogs of antibiotics, other than the penicillin and cephalosporins, have also been produced. The availability of truly efficacious antibiotics to treat a wide variety of severe infections undoubtedly represents one of the primary contributors to prolongation of life in modern society, as compared to the situation that existed in the early part of this century.

Coincidental with the above developments, biomedical scientists in pharmaceutical companies were actively pursuing purified extracts and pure compounds derived from plants and animal sources (e.g., digitalis, rauwolfia alkaloids, and animal hormones) as human medicaments. Analogs and derivatives of these purified substances were also investigated

intensively in the hope of increasing potency, decreasing toxicity, altering absorption, securing patent protection, etc. During this period, impressive discoveries were made in the fields of cardiovascular, central nervous system, and metabolic diseases (especially diabetes); medicinal chemists and pharmacologists set up programs to discover new and, hopefully, improved tranquilizers, antidepressants, antianxiety agents, antihypertensive agents, hormones, etc. Progress in the discovery of agents to treat cardiovascular and central nervous system diseases was considerably slower than was the case with infectious diseases. The primary reason for this delay is the relative simplicity and straightforwardness of dealing with an infectious disease as compared to diseases of the cardiovascular system or of the brain. Specifically, infectious diseases are caused by organisms that, in many cases, can be grown in test tubes, which markedly facilitates the rate at which compounds that inhibit the growth of, or actually kill, such organisms can be discovered. Not only was the testing quite simple when carried out in the test tube but also the amounts of compounds needed for laboratory evaluation were extremely small as compared to those required for animal evaluation. In addition, animal models of infectious diseases were developed very early in the history of this aspect of pharmaceutical research and activity in an intact animal as well as toxicity could be assessed in the early stages of drug discovery and development. Such was not the case in the 1950s as far as cardiovascular, mental, or certain other diseases were concerned because the basic defect or defects that lead to the disease in man were quite unknown. In addition, early studies had to be carried out in animal test systems, test systems which required considerable amounts of the compound and were much more difficult to quantitate than were the *in vitro* systems used in the infectious-disease field. The successes in the antibiotic field undoubtedly showed a carry-over or 'domino' effect in other areas of research as biochemists and biochemical pharmacologists began to search for *in vitro* test systems to provide more rapid screening for new drug candidates, at least in the cardiovascular and inflammation fields. The experimental dialog among biochemists, pharmacologists, and clinicians studying cardiovascular and mental diseases led, in the 1960s, to the development of various animal models of these diseases that increased the rate of discovery of therapeutic agents for the treatment thereof. Similar research activities in the fields of cancer research, viral infections, metabolic diseases, AIDS, inflammatory disease, and many others have, likewise, led to *in vitro* and animal models that have markedly increased the ability to discover new drugs in those important fields of research. With the increased discovery of drug activity came the need for increased regulation and, from the early 1950s on, the Food and Drug Administration (FDA) expanded its activities and enforcement of drug laws with both positive and negative results, from the standpoint of drug discovery.

In the later quarter of the 20th century, an exciting new technology emerged into the pharmaceutical scene, namely, biotechnology. Using highly sophisticated, biochemical genetic approaches, significant amounts of proteins, which, prior to the availability of so-called genetic engineering could not be prepared in meaningful quantities, became available for study and development as drugs. Furthermore, the new technology permitted scientists to isolate, prepare in quantity, and chemically analyze receptors in and on mammalian cells, which allows one to actually design specific effectors of these receptors.

As the drug discovery process increased in intensity in the mid- to late 20th century, primarily as a result of the major screening and chemical synthetic efforts in the pharmaceutical industry in industrialized countries worldwide, but also as a result of the biotechnology revolution, the need for increased sophistication and efficacy in (1) how to discover new drugs, (2) how to reproducibly prepare bulk chemicals, (3) how to determine the activity and safety of new drug candidates in preclinical animal models prior to their

administration to human beings and, finally, (4) how to establish their efficacy and safety in man, became of paramount importance. Likewise, the ability to reproducibly prepare extremely pure material from natural sources or biotechnology reactors on a large scale and to deliver stable and sophisticated pharmaceutical preparations to the pharmacists and physicians also became significant.

The above brief history of early drug use and discovery is intended to be purely illustrative and the reader is referred to an excellent treatise by Mann[1] to become well informed on the history of drug use and development from the earliest historic times to the present day.

Reference

1. Mann, R.D., *Modern Drug Use: An Enquiry on Historical Principles*, MTP Press, Lancaster, England, 1984, pp. 1–769.

Section I

General Overview

2

Overview of the Current Process of New Drug Discovery and Development

Charles G. Smith and James T. O'Donnell

CONTENTS

The first edition[1] of this book was published approximately 13 years ago. Its primary objective was to present an overview and a "roadmap" of the process of new drug discovery and development, particularly oriented to individuals or companies entering the pharmaceutical field. It was written by one of the authors (Smith), with no contributors, and drawn on Smith's experiences in the industry and field over the course of nearly 40 years. In the second edition, the scope of the first book has been expanded and technical details in the form of hard data have been included. In addition to the editors' own commentary and contributions, the major part of the book is the result of contributions of experts in the industry. New chapters on risk assessment, international harmonization of drug development and regulation, dietary supplements, patent law, and entrepreneurial startup of a new pharmaceutical company have been added. Some of the important, basic operational aspects of drug discovery and development (e.g., organizational matters, staff requirements, pilot plant operations, etc.) are not repeated in this book but can be found in the first edition.

In the 1990s and the new millennium, major changes have occurred in the pharmaceutical industry from the vantage points of research and development as well as commercial operations. New technologies and processes such as "high throughput screening" and "combinatorial chemistry" were widely embraced and developed to a high state of performance during this period. The very impressive rate of throughput testing the hundreds of thousands of compounds required micronization of operations, resulting in the reduction of screening reaction mixtures from milliliters to microliters. The systems are generally controlled by robots, and testing plates can accommodate a wide spectrum of biological tests. Combinatorial chemistry, a process in which a core molecule is modified with a broad spectrum of chemical reactions in single or multiple reaction vessels, can produce tens of thousands of compounds for screening. The objective of both approaches is to provide very large numbers of new chemical entities to be screened for biological activity *in vitro*. The use of computers to design new drug candidates has been developed to a significant level of sophistication. By viewing on the computer, the "active site" to which one wants the drug

candidate to bind, a molecule can often be designed to accomplish that goal. The true impact of these approaches on the actual rate of discovering new drugs is yet to be established. Some have questioned the utility of these new screening methods, claiming that no new molecular entities (NME) have resulted from these new screening methodologies, despite hundreds of millions invested by the industry.

Studies in the last few years in the fields of genomics and proteomics have made available to us an unprecedented number of targets with which to search for new drug candidates. While knowledge of a particular gene sequence, for example, may not directly point to a specific disease when the sequences are first determined, investigations of their presence in normal and diseased tissues could well lead to a quantitative *in vitro* test system that is not available today. The same can be said for the field of proteomics, but final decisions on the value of these new technologies cannot be made for some years to come.

Thanks to advances in genomics, animal models can now be derived using gene manipulation and cloning methods that give us never-before available *in vivo* models to be used in new drug screening and development. A large number of mammalian cell culture systems have also been developed not only to be used in primary screening but also for secondary evaluations. For example, the *in vitro* Caco 2 system shows some very interesting correlation with drug absorption *in vivo*. A test such as this is mandatory when one is dealing with several thousands of compounds or mixtures in a given experiment. More time will be needed to be absolutely certain of the predictability of such test systems but, appropriately, Caco 2 is widely used today in screening and prioritizing new drug candidates. As is always the case, the ultimate predictability of all the *in vitro* tests must await extensive studies in humans, which will occur several years henceforth.

In addition to the discussion are metabonomics that relate to their unique position within the hierarchy of cell function and their propensity to cross membranes and organs. Thus, many metabolites are found in bodily fluids that are accessible to measurement in humans using relatively noninvasive technologies. The study of metabolomics provides the pragmatic link from the macromolecular events of genomics and proteomics to those events recognized in histology. Applications of such strategies can potentially translate discovery and preclinical development to those metabolites measured traditionally, as first-in-human studies are performed earlier in drug discovery and development process, especially where no animal models are adequate.

During the past decade, clinical trial methodology has been expanded, improved, and, in large measure, standardized. The clinical testing phase of new drug development is the most expensive single activity performed. In addition to cost, it is very time consuming since, with chronic diseases, one must investigate the new drug candidate in a significant number of patients over a period of months or years, in randomized, double-blind, placebo- or active-drug-controlled studies. The search for surrogate endpoints continues, as it should, because a surrogate endpoint can markedly increase the rate of progression in clinical investigations with new drug candidates in certain disease states. Modern advances in molecular biology, receptor systems, cellular communication mechanisms, genomics, and proteomics will, according to our belief, provide researchers with new approaches to the treatment of a variety of chronic diseases. Significantly improved prescription medications are sorely needed in many fields. In the past decade, we have witnessed very impressive advances in the treatment of AIDS, for example. There is no question that life expectancy has been increased, albeit accompanied by significant drug toxicity and the need to use a "cocktail" of drugs in combination. The ability of the AIDS virus to mutate and become drug resistant presents a major and imminent threat to all patients afflicted with this disease. Serious efforts are under way in the pharmaceutical industry to find new drugs, across the entire infections diseases spectrum, which are not cross-resistant with existing therapies.

Cancer and AIDS vaccines are also under investigation using new technologies and, hopefully, the day will come when we can prevent or ameliorate some of these debilitating and fatal diseases by vaccination. In the cancer field, new methodologies in science have, again, given us new targets with which to search for chemotherapeutic agents. The humanization of monoclonal antibodies has resulted in the marketing of some truly impressive drugs that are much better tolerated by the patient than are cytotoxic agents. In the case of certain drug targets in cancer, impressive results have been seen in the percentage of leukemia and lymphoma patients who can be brought into complete remission. In addition, biological medications to increase red and white blood cells have become available. Unfortunately, drug resistance once again plagues the cancer field, as are the cases with AIDS and various infectious diseases. As a result, researchers are seeking compounds that are not cross-resistant with existing therapies. Very significant advances in drug discovery are also expected to be seen in central nervous system, cardiovascular, and other chronic diseases as a result of breakthrough research in these fields.

Although the focus of this book is the research and development side of the pharmaceutical industry, certain commercial considerations are worth mentioning because of the major impact they may have on new drug research. These opinions and conclusions are based solely on decades of experience in the field by editors, working in the industry within companies and as an independent consultant (Smith), and also as a health care worker and academic (O'Donnell). No financial incentive for these statements has been received from the pharmaceutical industry. As the result of the very complicated nature of drug discovery and development, unbelievable costs accrue in order to bring a new therapeutic agent to market. Increasing costs are incurred, in part, from (1) shifting disease targets from more rapidly evaluable, acute diseases to those with poor endpoints and chronicity and (2) the emergence and rapid spread of serious diseases in society (e.g., AIDS, certain cancers, hepatitis C, etc.). In addition to increasing cost, the time required to gather sufficient data to be able to prove, to a statistically valid endpoint, that the drug has indeed been effective in a given disease has risen. The cost for the development of a major drug has been widely stated to be US $800 million per new therapeutic agent placed on the market.[1] This figure incorporates, of course, the cost of "lost" compounds that did not make the grade during preclinical or clinical testing. It has recently been reported that, while historically 14% of drugs that entered phase I clinical trials eventually won approval, now only 8% succeed. Furthermore, 50% of the drug candidates fail in the late stage of phase III trials compared to 20% in past years. More details on these points can be found in the literature (cf., Refs. 2–8).

The average time from the point of identifying a clinical candidate to approval of a new drug is approximately 10 years. There is an understandable clamor in the population and in our legislative bodies to lower the price of prescription drugs. The cost of some prescription drugs is, to be sure, a serious problem that must be addressed but some of the solutions, suggested and embraced by certain legislators, could have serious negative impact on new drug discovery and development in the future. For example, allowing the importation of prescription drugs from Canada or other non-U.S. countries (25 around the world have been mentioned) may well reduce the price of new drugs in this country to the point of significantly decreasing profits that are needed to support the tremendous cost of new drug discovery and development. The record clearly shows that countries that control drug prices, frequently under socialist governments, do not discover and develop new prescription drugs. The reason is obvious since the cost and time factors for new drug discovery can only be borne in countries in which the pharmaceutical companies are clearly profitable. Our patent system and lack of price controls are the primary reasons for the huge industrial success of new product development in this country, in and out of the pharmaceutical arena. If we undercut that system in the prescription drug field, the cost

of drugs will certainly go down in the United States in the short term but, without the necessary profits to invest heavily in new drug discovery and development, the latter will also surely drop. Since it requires a decade from the time of initial investigation to marketing of a new drug, this effect would not be evident immediately after (1) allowing reimportation, (2) overriding patent protection, or (3) implementing price controls but, within a period of 5 to 10 years, we would certainly see pipelines of new medications beginning to dry up. Indeed, if such a system were allowed to continue for several years, new drug development as we know it would, in our opinion, be seriously impeded. When legislators look to Canada as an example of successful government subsidy of drugs, they should also consider whether a country like Canada could ever produce a steady stream of major new drugs, as does the United States. Research budgets have never been larger, we have never had as many innovative and exciting targets on which to focus, and this enormous effort cannot be afforded unless the companies selling the drugs can realize an adequate profit. If our pipelines of new prescription drugs dry up, you can be rest assured that the deficit will not be satisfied elsewhere in the world. It has been reported that, 10 years ago drug companies in Europe produced a significantly greater percentage of prescription drugs than is the case today. Society simply cannot afford to risk a marked reduction in new drug discovery in this country. Patients must join the fight to see that activities to impose price controls, which will inevitably reduce the rate of discovery of many potential drugs, are not based on political motives on the part of legislators. At this point in history, U.S. science stands in the forefront of new drug discovery and development. As noted above, never before have we had such an array of biological targets and synthetic and biotechnological methods with which to seek new medications. Hopefully, our government, in collaboration with the pharmaceutical industry, will find more suitable methods to solve the question of the cost of new pharmaceuticals than to impose price controls equal to those in countries that have socialized medicine. There can be no question as to whether the primary loser in such moves will be patients.

In addition to the question of the rate of drug discovery and development, we must be concerned about the quality of drugs available by mail or over the internet. The Food and Drug Administration (FDA) cannot possibly afford to check all drugs flowing into America from as many as 25 foreign countries from which our citizens might be allowed to buy prescription drugs. It will be interesting to compare the regulatory requirements for FDA approval in the United States with those of the least stringent of the foreign countries from which some of our legislators want to approve importation of drugs. Would Congress be prepared to mandate a lowering of FDA standards to the same level in order to reduce the cost of drug discovery and development in this country? We certainly hope not! Indeed, there have been reports that drugs imported and sold on the internet are counterfeit, and frequently contain little or no labeled active ingredients, and further, may contain adulterants.

Another new topic chapter in the second edition of this book discusses the so-called dietary supplements, contributed by a recognized authority in Health Fraud. Over the past few years and, especially, since the passage of the DSHEA Act by Congress,[9] the use of such products has increased dramatically and they are made widely available to the public with little or no FDA regulation. Although the law prevents manufacturers from making a medical treatment claim on the label of these preparations, such products generally have accompanying literature citing a variety of salutary effects in patients with various ills, the majority of which have not been proven by FDA type-randomized, double-blind, placebo-controlled clinical studies, of the kind that must be performed on prescription drugs and some "over-the-counter" drugs in this country. Published studies on quality control defects in some of these dietary supplement products (cf. ConsumerLab.com) indicate the need for tightening up of this aspect of product development. FDA is currently promulgating GMPs for dietary supplements. An enhanced

enforcement of the dietary supplement regulations now exists.[9] A small segment of the dietary supplement industry has been calling for GMPs and increased FDA regulation.[10]

2.1 Basic Scientific Discovery and Application to New Drug Development

In an apparent attempt to determine whether the American taxpayer is getting fair benefits from research sponsored by the federal government, the Joint Economic Committee of the U.S. Senate (for history see Ref. 7) has been considering this question. Historically, basic research has been funded by the NIH and various philanthropic foundations to discover new concepts and mechanisms of bodily function, in addition to training scientists. The role of industry has been to apply the basic research findings to specific treatments or prevention of disease. This is the appropriate manner in which to proceed. The industry cannot afford to conduct sufficient basic research on new complicated biological processes in addition to discovering new drugs or vaccines. The government does not have the money, time, or required number of experts to discover and develop new drugs.

The process that plays out in real life involves the focus of pharmaceutical industry scientists on desirable biological targets that can be identified in disease states, and to set up the program to discover specific treatments that will show efficacy in human disease. The compounds that are developed successfully become drugs on which the company holds patents. In this manner, the enormous cost of discovering and developing a new drug (estimated at $800 million plus over a period of some 10 years[1]) as noted above can be recouped by the founding company since no competitors can sell the product as long as the patent is in force. Without such a system in place, drug companies simply could not, in our opinion, afford to bring new prescription drugs to the market.

In the course of reviewing the matter, the Joint Economic Committee examined a list of 21 major drugs, which was put together apparently as an example of drug products that might justify royalty to the government. One of these agents, captopril (trade name Capoten), was discovered and developed by E.R. Squibb & Sons in the 1970s. At that time, Charles Smith (one of the authors/editors) was vice president for R&D at The Squibb Institute for Medical Research. One of Squibb's academic consultants, Professor Sir John Vane of the Royal College of Surgeons in London brought the idea of opening a new pathway to treat the so-called essential hypertension by inhibiting an enzyme known as the angiotensin converting enzyme (ACE). This biochemical system was certainly known at that time but, in Squibb's experience in the field of hypertension treatment, was not generally thought to play a major role in the common form of the disease, then known as "essential hypertension." The company decided to gamble on finding a treatment that was not used at the time and that would be proprietary to the company. Professor Vane (Nobel laureate in medicine in 1982) had discovered a peptide in snake venom that was a potent inhibitor of ACE. Squibb decided to pursue the approach he espoused, resulting in the development of a unique treatment for this very prevalent and serious disease.

In the first phase of their research, Squibb tested a short-chain peptide isolated from the venom of the viper *Bothrops jararaca*, with which Vane was working in the laboratory, in human volunteers and showed that it did, indeed, inhibit the conversion of angiotensin I to angiotensin II after intravenous injection. The peptide was also shown to reduce blood pressure in patients when injected. Since the vast majority of peptides cannot be absorbed from the GI tract, Squibb scientists set out to prepare a nonpeptide compound that could be used orally and manufactured at acceptable cost. The design of a true peptidomimetic that became orally active had not been accomplished at that time. Squibb then carried out

a full-blown clinical program on a worldwide basis, which led to FDA approval of Squibb's drug Capoten (captopril), an ACE inhibitor. Mark also marketed an ACE inhibitor in the same time frame. This work opened a new area of research that has resulted in a bevy of new drugs that share this mechanism of action for use as antihypertensive drugs (for more detail, see Refs. 11–15).

In the minds of pharmaceutical researchers and, hopefully, the public at large, the above example illustrates the unique role of pharmaceutical companies in making good use of basic research to discover new treatments for serious diseases. The huge costs to discover and develop a new drug could not be borne unless the companies knew that, if their gamble worked (which is not the case in the majority of situations), they would be assured of a good financial return for their shareholders. This system has served the country well in many fields of endeavor, in and out of the drug arena, and should be retained as such.

2.2 Regulation of New Drug Development

Drug development will come to a crashing halt without approval of the U.S. FDA, authorized by Congress to approve, license, and monitor the drugs sold to the American public. We are fortunate to have two contributors from the FDA, an acting associate commissioner for operations, and also CDER's (Center for Drug Evaluation and Research) associate director for International Conference on Harmonisation (ICH). These authors describe the FDA's new critical pathway initiative, pharmacists' risk management contributions, as well as the Common Technical Document (eCTD), which will enable a sponsor to file in one of the cooperating ICH partners, and receive approval for almost global marketing of the new agent. A very important chapter on pharmacogenetics and pharmacogenomics includes numerous FDA contributers.

2.3 Liability and Litigation

Last and the most unpopular topic in any industry, especially in the pharmaceutical industry, is the topic of liability and litigation. We have elected to include a chapter on this topic so that workers from all scientific disciplines involved in drug discovery and development can learn from history, and, hopefully, avoid being involved in the devastation of life (due to toxicity of inadequately manufactured drugs or drugs with inadequate warnings for safe use) and destruction of companies and careers that follows in the aftermath of drug product litigation.

References

1. Smith, C.G., *The Process of New Drug Discovery and Development*, 1st ed., CRC Press, Boca Raton, FL, 2002.
2. Di Masi, J.A., Hansen, R.W., and Grabowski, H.G., The price of innovation: new estimates of drug development costs, *J. Health Econ.*, 22, 151–185, 2003.

3. Reichert, J.M. and Milne, C.-P., Public and private sector contributions to the discovery and development of 'impact' drugs, *Am. J. Therapeut.*, 9, 543–555, 2002.
4. Di Masi, J., Risks in new drug development. Approval success rates for investigational drugs, *Clin. Pharmacol. Therapeut.*, 69, 297–307, 2001.
5. Di Masi, J., New drug development in the United States from 1963 to 1999, *Clin. Pharmacol. Therapeut.*, 69, 286–296, 2001.
6. Grabowski, H., Vernon, J., and Di Masi, J.A., Returns on research and development for 1990s new drug introductions, *Pharmacol. Econ.*, 20 (suppl. 3), 11–29, 2002.
7. Reichert, J.M. and Milne, C.-P., Public and private sector contributions to the discovery and development of "impact" drugs, A Tufts Center for the Study of Drug Development White Paper, May 2002.
8. Hardin, A., More compounds failing phase I, *Scientist Daily News*, Aug. 6, 2004, p. 1.
9. Dietary Supplement Health and Education Act of 1994, Publ. no. 103-417, 108 Stat. 4325 codified 21 U.S.C. 321, et seq. (suppl. 1999).
10. FDA links: (a) http://www.cfsan.fda.gov/~dms/ds-warn.html (b) http://www.cfsan.fda.gov/~lrd/ hhschomp.html (c) http://www.fda.gov/ola/2004/dssa0608.html
11. Smith, C.G. and Vane, J.R., The discovery of captopril, *FASEB J.*, 17, 788–789, 2003.
12. Gavras, H., The discovery of captopril: Reply, *FASEB J.*, 18, 225, 2004.
13. Erdas, E.G., The discovery of captopril: Reply, *FASEB J.*, 18, 226, 2004.
14. Pattac, M., From viper's venom to drug design: treating hypertension, *FASEB J.*, 18, 421, 2004.
15. Smith, C.G. and Vane, J.R., The discovery of captopril: Reply, *FASEB J.*, 18, 935, 2004.

3

Integrated Drug Product Development — From Lead Candidate Selection to Life-Cycle Management

Madhu Pudipeddi, Abu T.M. Serajuddin, and Daniel Mufson

CONTENTS

3.1 Introduction

Historically, medicines have been administered through the obvious portals following their preparation first by the shaman and then by the physician and later by the apothecary. These natural products were ingested, rubbed-in, or smoked. For the past century, the person diagnosing the disease no longer prepares the potion, eliminating, no doubt, some of the power of the placebo, and as a consequence, drug discovery, development, and manufacturing have grown into a separate pharmaceutical industry. In particular, the last 50 years have been a period of astounding growth in our insight of the molecular function of the human body. This has led to discovery of medicines to treat diseases that were not even recognized a half-century ago. This chapter reflects the role of pharmaceutics and the diversity of the approaches taken to achieve these successes, including approaches that were introduced within recent years, and describes how the role of the "industrial" pharmacist has evolved to become the technical bridge between discovery and development activities and, indeed, commercialization activities. No other discipline follows the progress of the new drug candidate as far with regard to the initial refinement of the chemical lead through preformulation evaluation to dosage-form design, clinical trial material (CTM) preparation, process scale-up, manufacturing, and then life-cycle management (LCM).

The pharmaceutical formulation was once solely the responsibility of the pharmacist, first in the drugstore and later in an industrial setting. Indeed, many of today's major drug companies, such as Merck, Lilly, Wyeth, and Pfizer components Searle, Warner-Lambert, and Parke-Davis, started in the backrooms of drugstores. During the second half of the 20th century, physicochemical and biopharmaceutical principles underlying pharmaceutical dosage forms were identified and refined, thanks to the pioneering works by Higuchi,[1] Nelson,[2] Levy,[3] Gibaldi,[4] and their coworkers. Wagner,[5] Wood,[6] and Kaplan[7] were among the earliest industrial scientists to systematically link formulation design activities and biology. Nevertheless, until recently, formulations were developed somewhat in isolation with different disciplines involved in drug development operating independently. For example, during the identification and selection of new chemical entities (NCEs) for development, not much thought was given into how they would be formulated, and during dosage-form design, adequate considerations of *in vivo* performance of formulations was lacking. Wagner[5] first termed our evolving understanding of the relationship between the dosage form and its anatomical target, "biopharmaceutics" in the early 1960s. Since then it has been apparent that careful consideration of a molecule's physical chemical properties and those of its carrier, the dosage form, must be understood to enhance bioavailability, if given orally, and to enhance the ability of drug to reach the desired site of action, if given by other routes of administration. This knowledge allows for a rational stepwise approach in selecting new drug candidates, developing

optimal dosage forms, and, later when it is necessary, making changes in the formulation or manufacturing processes. During the last decade or so, the basic approach of dosage-form development in the pharmaceutical industry has changed dramatically. Dosage-form design is now an "integrated process" starting from identification of drug molecules for development to their ultimate commercialization as dosage forms. This is often performed by a multidisciplinary team consisting of pharmacists, medicinal chemists, physical chemists, analytical chemists, material scientists, pharmacokineticists, chemical engineers, and other individuals from related disciplines.

In its simplest terms the dosage form is a carrier of the drug. It must further be reproducible, bioavailable, stable, readily scaleable, and elegant. The skill sets employed to design the first units of a dosage form, for example, a tablet, are quite different than those required to design a process to make hundreds of thousands of such units per hour, reproducibly, in ton quantities, almost anywhere in the world. Nevertheless, it is important that "design for manufacturability" considerations are made early although resource constraints and minimal bulk drug supply may not favor them. The manufacturability situation becomes understandably more complex as the dosage form becomes more sophisticated or if a drug-delivery system (DDS) is needed.

The level of sophistication in dosage-form design has been keeping pace with advances in discovery methods. New excipients, new materials, and combination products that consist of both a drug and a device have arisen to meet new delivery challenges. For example, many of the NCEs generated by high-throughput screening (HTS) are profoundly water-insoluble. What was considered a lower limit for adequate water solubility[7] (~0.1 mg/mL) in the 1970s has been surpassed by at least an order of magnitude due to changes in the way drug discovery is performed. Traditional methods such as particle size reduction to improve the aqueous dissolution rate of these ever more insoluble molecules are not always sufficient to overcome the liability. New approaches have evolved to meet these challenges ranging from cosolvent systems[8] to the use of lipid—water-dispersible excipients[9] and to the establishment of numerous companies with proprietary methods to increase bioavailability.

Many literature sources describing formulation and manufacture of different pharmaceutical dosage forms are available.[10,11] The primary objective of this chapter is to describe an integrated process of drug development, demonstrating how all activities from lead selection to LCM are interrelated. Various scientific principles underlying these activities are described.

A survey of new drug approvals (NDAs) during the last 5 years (1999 to mid-2004) showed that nearly 50% of them are oral dosage forms. The percentage is higher if Abbreviated NDAs for generics are included. Therefore, the primary focus of this chapter is the development of oral dosage forms with a few other dosage forms described only briefly. However, many of the principles described in this chapter are common to all dosage forms.

3.2 Developability Assessment

The dosage-form design is guided by the properties of the drug candidate. If an NCE does not have suitable physical and chemical properties or pharmacokinetic attributes, the development of a dosage form (product) may be difficult and may sometimes be even impossible. Any heroic measures to resolve issues related to physicochemical and biopharmaceutical properties of drug candidates add to the time and cost of drug

development. Therefore, in recent years, the interaction between discovery and development scientists increased greatly to maximize the opportunity to succeed.[12]

3.2.1 Evolution of the Drug Discovery and Development Interaction

The traditional (i.e., pre-1990s) drug discovery process involved initial lead generation on the basis of natural ligands, existing drugs, and literature leads. New compounds would be synthesized and tested for biological activity, and structure–activity relationships would be established for optimization of leads using traditional medicinal chemistry techniques. Promising compounds would then be promoted for preclinical and clinical testing and therefore passed along to the product development staff. While often there was little collaboration between research and development, a few organizations had recognized the importance of discovery-development teams to assess development issues related to new drug candidates.[7]

The current (post-1990s) drug discovery process typically involves:[13]

- Target identification
- Target validation
- Lead identification
- Candidate(s) selection

A drug target can be a receptor/ion channel, enzyme, hormone/factor, DNA, RNA, nuclear receptor, or other, unidentified, biological entity. Once drug targets are identified, they are exposed to a large number of compounds in an *in vitro* or cell-based assay in an HTS mode. Compounds that elicit a positive response in a particular assay are called "hits." Hits that continue to show positive response in more complex models rise to "leads" (lead identification). A selected few of the optimized leads are then advanced to preclinical testing. The traditional discovery process has not been discontinued but still occurs in a semiempirical fashion depending on the chemist's or biologist's experience and intuition. With the application of HTS technologies, compound handling in discovery has shifted to the use of organic stock solutions (dimethylsulfoxide) for *in vitro* and *in vivo* testing from the traditional use of gum tragacanth suspensions in rats by the pharmacologist.

Use of combinatorial chemistry and HTS technologies have resulted in the generation and selection of increasingly lipophilic drug molecules with potential biopharmaceutical hurdles in downstream development.[14] Particularly, the use of organic solvents such as dimethylsulfoxide has contributed to the increase in water-insoluble drugs. Analysis of compound attrition in pharmaceutical development indicated that poor pharmacokinetic factors, i.e., absorption, elimination, distribution, and metabolism (ADME) contributed to about 40% of failed candidates, and for those that moved forward, the development timelines significantly slowed down.[15] To reduce attrition of compounds later in development, pharmaceutical companies began to conduct pharmaceutical, pharmacokinetic, and safety profiling of late- as well as early-phase discovery compounds.[16]

3.2.2 Screening for Drugability or Developability

Compounds with acceptable pharmaceutical properties, in addition to acceptable biological activity and safety profile, are considered "drug-like" or developable. Typical acceptable pharmaceutical properties for oral delivery of a drug-like molecule include sufficient aqueous solubility, permeability across biological membranes, satisfactory stability to

metabolic enzymes, resistance to degradation in the gastrointestinal (GI) tract (pH and enzymatic stability), and adequate chemical stability for successful formulation into a stable dosage form. A number of additional barriers, such as efflux transporters[17] (i.e., export of drug from blood back to the gut) and first-pass metabolism by intestinal or liver cells, have been identified that may limit oral absorption. A number of computational and experimental methods are emerging for testing (or profiling) drug discovery compounds for acceptable pharmaceutical properties.

In this section, discussion of physicochemical profiling is limited to solubility, permeability, drug stability, and limited solid-state characterization (as we will see in Section 3.4, there are other physical–mechanical properties that must also be considered). For convenience, methods available for physicochemical profiling are discussed under the following categories: computational tools (sometimes referred to as *in silico* tools), HTS methods, and in-depth physicochemical profiling.[16]

3.2.2.1 Computational Tools

Medicinal chemists have always been adept in recognizing trends in physicochemical properties of molecules and relating them to molecular structure. With rapid increase in the number of hits and leads, computational tools have been proposed to calculate molecular properties that may predict potential absorption hurdles. For example, Lipinski's "Rule of 5"[14] states that poor absorption or permeation are likely when:

1. There are more than five H-bond donors (expressed as the sum of –NH and –OH groups).
2. The molecular weight is more than 500.
3. log P>5 (or c log P>4.5).
4. There are more than ten H-bond acceptors (expressed as the sum of Ns and Os)

If a compound violates more than two of the four criteria, it is likely to encounter oral absorption issues. Compounds that are substrates for biological transporters and peptidomimetics are exempt from these rules. The Rule of 5 is a very useful computational tool for highlighting compounds with potential oral absorption issues. A number of additional reports on pharmaceutical profiling and developability of discovery compounds have been published,[18] since the report of Rule of 5. Polar surface area (PSA) and number of rotatable bonds have also been suggested as means to predict oral bioavailability. PSA is defined as the sum of surfaces of polar atoms in a molecule. A rotatable bond is defined as any single bond, not in a ring, bound to a nonterminal heavy (i.e., non-hydrogen) atom. Amide bonds are excluded from the count. It has been reported that molecules with the following characteristics will have acceptable oral bioavailability:[19]

1. Ten or fewer rotatable bonds.
2. Polar surface area equal to or less than 140 Å^2 (or 12 or fewer H-bond donors and acceptors).

Aqueous solubility is probably the single most important biopharmaceutical property that pharmaceutical scientists are concerned with. It has been the subject of computational prediction for several years.[20–23] The overall accuracy of the predicted values can be expected to be in the vicinity of 0.5 to 1.0 log units (a factor of 3 to 10) at best. Although a decision on acceptance or rejection of a particular compound cannot be made only on the basis of predicted parameters, these predictions may be helpful to direct chemical libraries with improved drug-like properties.[24]

3.2.2.2 High-Throughput Screening Methods

High-throughput drug-like property profiling is increasingly used during lead identification and candidate selection. HTS pharmaceutical profiling may include:

- Compound purity or integrity testing using methods such as UV absorbance, evaporative light scattering, MS, NMR, etc.[25]
- Solubility.
- Lipophilicity (log *P*).
- Dissociation constant (pK_a).
- Permeability.
- Solution/solid-state stability determination.

Compound purity (or integrity testing) is important to ensure purity in the early stages because erroneous activity or toxicity results may be obtained by impure compounds. It is initiated during hit identification and continued into lead and candidate selection.

Solubility is measured to varying degrees of accuracy by HTS methods. Typical methods in the lead identification stage include determination of "kinetic solubility" by precipitation of a drug solution in dimethylsulfoxide into the test medium. Since the solid-state form of the precipitate (crystalline or amorphous) is often not clearly known by this method, the measured solubility is approximate and generally higher than the true (equilibrium) solubility. Kinetic solubility, however, serves the purpose of identifying solubility limitations in activity or *in vitro* toxicity assays or in identifying highly insoluble compounds. Lipinski et al.[14] observed that, for compounds with a kinetic solubility greater than 65 μg/mL (in pH 7 non-chloride containing phosphate buffer at room temperature), poor oral absorption is usually due to factors unrelated to solubility. The acceptable solubility for a drug compound depends on its permeability and dose. This point will be further elaborated later. Methods to improve solubility in lead optimization have been reviewed.[26]

Estimation or measurement of pK_a is important to understand the state of ionization of the drug under physiological conditions and to evaluate salt-forming ability.[27] Log *P* determines the partitioning of a drug between an aqueous phase and a lipid phase (i.e., lipid bilayer). Log *P* and acid pK_a can be theoretically estimated with reasonable accuracy.[14,28,29] High-throughput methods are also available for measurement of log *P*[30] and pK_a.[31]

Physical flux of a drug molecule across a biological membrane depends on the product of concentration (which is limited by solubility) and permeability. High-throughput artificial membrane permeability (also called Parallel Artificial Membrane Permeability Assay) has been used in early discovery to estimate compound permeability.[32] This method measures the flux of a compound in solution across an artificial lipid bilayer deposited on a microfilter. Artificial membrane permeability is a measure of the actual flux (rate) across an artificial membrane whereas log *P* or log *D* — as mentioned earlier — represent equilibrium distribution between an aqueous and a lipid phase. Sometimes the term "intrinsic permeability" is used to specify the permeability of the unionized form. Artificial membrane permeability can be determined as a function of pH. The fluxes across the artificial membrane in the absence of active transport have been reported to relate to human absorption through a hyperbolic curve. The correlation of permeability through artificial membranes may depend on the specific experimental conditions such as the preparation of the membranes and pH. Therefore, guidelines on what is considered acceptable or unacceptable permeability must be based on the individual assay conditions. For example, Hwang et al.[33] ranked compound permeation on the basis of the percent transport across the lipid bilayer in 2 h: <2% (low), 2 to 5% (medium), and >5% (high), respectively.

Caco-2 monolayer, a model for human intestinal permeability, is commonly used in drug discovery to screen discovery compounds.[34,35] The method involves measurement of flux of

the compound dissolved in a physiological buffer through a monolayer of human colonic cells deposited on a filter. Caco-2 monolayer permeability has gained considerable acceptance to assess human absorption. Compounds with a Caco-2 monolayer permeability (P_{app}) similar to or greater than that of propranolol (~30×10^{-6} cm/sec) are considered highly permeable, while compounds with P_{app} similar to or lower than that of ranitidine (<1×10^{-6} cm/sec) are considered poorly permeable. Hurdles associated with determination of permeability of poorly soluble compounds using Caco-2 method have been reviewed.[36]

3.2.2.3 In-Depth Physicochemical Profiling

Once compounds enter the late lead selection or candidate selection phase, more in-depth physicochemical profiling is conducted. The extent of characterization may vary from company to company; however, it likely includes:

- Experimental pK_a and log P (as a function of pH, if necessary)
- Thermodynamic solubility (as a function of pH)
- Solution/suspension stability
- Solid-state characterization

Solid-state characterization typically involves:

- Solid-state stability
- Feasibility of salt formation
- Polymorph characterization
- Particle size, hygroscopicity
- Dissolution rate

In a more traditional pharmaceutical setting, this characterization would be done during preformulation studies. With the availability of automation and the ability to conduct most of these experiments with small quantities of material, more preformulation activities are being shifted earlier into drug discovery. Recently, Balbach and Korn[37] reported a "100 mg approach" to pharmaceutical evaluation of early development compounds. Additional absorption, metabolism, distribution, elimination, and toxicity[38] screens may also be conducted at this stage.

Overall, the scientific merit of physicochemical profiling is clear. It provides a better assessment of development risks of a compound early on. The important question is how can pharmaceutical companies utilize the vast amount of physicochemical information to advance the right drug candidates to preclinical and clinical testing? Scorecards or flags may be used to rank drug candidates for their physicochemical properties. These scores or flags, however, have to be appropriately weighted with biological activity, safety, and pharmacokinetic profiling of compounds. The relative weighting of various factors depends on the specific issues of a discovery program. However, the basic question of how does it help reduce attrition due to unacceptable physicochemical properties remains to be answered in a statistical sense. In 1997, Lipinski et al.[14] reported that a trend had been seen since the implementation of the Rule of 5 toward more drug-like properties in Pfizer's internal drug base. Overall, the goal of a discovery program is to steer the leads in the right direction using computational and HTS approaches and then utilize the in-depth screening tools to select the most optimal compound without undue emphasis on a single parameter such as biological activity. The overall success of a compound is a function of its biological and biopharmaceutical properties.

3.3 Overview of Dosage-Form Development and Process Scale-Up

Once a compound is selected for development, the pharmaceutics group begins a series of studies to further evaluate the salient physical-chemical and physical-mechanical properties of NCEs to guide the actual design (formulation, recipe) of the dosage form that will carry the drug. These investigations consist of the following steps that span the years that the molecule undergoes clinical evaluation:

- Preformulation
- Consideration of its biopharmaceutical aspects
- Dosage-form design
- CTMs manufacture
- Scale-up studies including technical transfer to manufacturing sites
- Initiation of long-term stability studies to guide the setting of the expiration date
- Production of "biobatches"
- Validation and commercial batches
- Life-cycle management

While there is a natural desire to front-load these evaluations, this must be balanced against the sad fact that many molecules fail to survive clinical testing: sufficient characterization is performed to help select the "right" molecule to minimize losses at the later, and much more costly, clinical evaluation stages. Great thought and planning are required to optimize the level of effort expended on a single molecule when so many are known to fail during clinical evaluation.

There are many reference on the design[10,11] and scale-up[39] of pharmaceutical dosage forms. It is appropriate to mention that the design of the dosage form be well documented from its inception as this information is required at the NDA stage to explain the development approach. The FDA and its global counterparts are seeking documentation that the product quality and performance are achieved and assured by design of effective and efficient manufacturing processes. The product specifications should be based upon a mechanistic understanding of how the formulation and processing factors impact the product performance. As we describe later, it is important that an ability to effect continuous improvement to the production process with continuous real-time assessment of quality must be incorporated in drug development. Recent regulatory initiatives require that the product/process risks are assessed and mitigated. In this respect, GMP regulations for drugs are moving toward quality standards already required for devices.

3.4 Preformulation

In the pre-1990s development scenario, preformulation activities would start when a compound had been chosen as the lead candidate by the drug discovery unit and advanced for preclinical development prior to testing in man. In the current integrated discovery-development scenario, many of the classical preformulation activities are conducted while screening compounds are in the lead identification or compound selection stage. Irrespective of the actual timing of preformulation studies, preformulation lays the foundation for robust formulation and process development. The purpose of preformulation is to understand the basic physicochemical properties of the drug compound so that the

challenges in formulation are foreseen and appropriate strategies are designed. Some physicochemical properties are independent of the physical form (i.e., crystal form) but simply are a function of the chemical nature of the compound. They include chemical structure, pK_a, partition coefficient, log P, and solution-state stability. A change in the physical form does not affect these properties.

3.4.1 Preformulation Activities: Independent of Solid Form

3.4.1.1 Dissociation Constant

The pK_a, or dissociation constant, is a measure of the strength of an acid or a base. The dissociation constant of an organic acid or base is defined by the following equilibria:

$$\begin{aligned} &HX + H_2O \rightleftharpoons H_3O^+ + X^- \text{ (acid)} \\ &BH^+ + H_2O \rightleftharpoons H_3O^+ + B \text{ (base)} \\ &K_a = \frac{[H_3O^+][X^-]}{[HX]} \quad \text{(acid)} \\ &K_a = \frac{[H_3O^+][B]}{[HB^+]} \quad \text{(base)} \end{aligned} \tag{3.1}$$

whereby the $-\log K_a$ is defined as pK_a.

The pK_a of a base is actually that of its conjugate acid. As the numeric value of the dissociation constant increases (i.e., pK_a decreases), the acid strength increases. Conversely, as the acid dissociation constant of a base (that of its conjugate acid) increases, the strength of the base decreases. For a more accurate definition of dissociation constants, each concentration term must be replaced by thermodynamic activity. In dilute solutions, concentration of each species is taken to be equal to activity. Activity-based dissociation constants are true equilibrium constants and depend only on temperature. Dissociation constants measured by spectroscopy are "concentration dissociation constants." Most pK_a values in the pharmaceutical literature are measured by ignoring activity effects and therefore are actually concentration dissociation constants or apparent dissociation constants. It is customary to report dissociation constant values at 25°C.

Drug dissociation constants are experimentally determined by manual or automated potentiometric titration or by spectrophotometric methods.[40] Current methods allow determination of pK_a values with drug concentrations as low as 10 to 100 μM. For highly insoluble compounds (concentration <1 to 10 μM), the Yesuda–Shedlovsky method[41] is commonly used where organic cosolvents (i.e., methanol) are employed to improve solubility. The method takes three or more titrations at different cosolvent concentrations, and the result is then extrapolated to pure aqueous system. The dissociation constant can also be determined with less accuracy from the pH–solubility profile using the following modification of Henderson–Hasselbach equation:

$$S = \begin{cases} S_{HA}(1 + 10^{pH - pK_a}) & \text{for an acid} \\ S_B(1 + 10^{pK_a - pH}) & \text{for a base} \end{cases} \tag{3.2}$$

where S_{HA} or S_B is the intrinsic solubility of the unionized form.

Some drugs exhibit concentration-dependent self-association in aqueous solutions. The dissociation constant of these compounds may change upon self-association. For example,

pK_a of dexverapamil has been reported to shift from 8.90 to 7.99 in the micellar concentration range.[42]

3.4.1.2 Partition or Distribution Coefficient

The partition coefficient (P) is a measure of how a drug partitions between a water-immiscible lipid (or an organic phase) and water. It is defined as follows:

$$P = \frac{[\text{neutral species}]_o}{[\text{neutral species}]_w} \qquad \log P = \log_{10} P \tag{3.3}$$

The distribution coefficient is the partition coefficient at a particular pH. The following equilibrium is often used to define D, with the assumption that only the unionized species partition into the oil or lipid phase:

$$D = \frac{[\text{unionized species}]_o}{[\text{unionized species}]_w + [\text{ionized species}]_w}$$

$$\log D_{\text{at pH}} = \begin{cases} \log P - \log\left[1 + 10^{(\text{pH}-\text{p}K)}\right] & \text{for acids} \\ \log P - \log\left[1 + 10^{(\text{p}K-\text{pH})}\right] & \text{for bases} \end{cases} \tag{3.4}$$

The measurement of log P is important because it has been shown to correlate with biological activity and toxicity.[43] As discussed in the previous section, a range of log P (0 to 5) has been shown to be critical for satisfactory oral absorption of drug compounds.

3.4.1.3 Solution Stability Studies

Forced degradation studies provide information on drug degradation pathways, potential identification of degradation products in the drug product, structure determination of degradation products, and determination of intrinsic stability of a drug. Regulatory guidance and best practices for conducting forced degradation studies have been reviewed.[44] Typical conditions for forced degradation testing include strong acid/base, oxidative, photostability, thermal, and thermal/humidity conditions. Stress conditions are utilized that result in approximately 10% degradation.

Stability of the drug in solution over a pH range of 1 to 13 is assessed during preformulation. The purpose is twofold. In the short run, stability in the GI pH range of 1 to 7.5 is important for drug absorption. In the long run, knowledge of solution-state stability is important for overall drug product stability and possible stabilization strategy. Drug degradation in solution typically involves hydrolysis, oxidation, racemization, or photodegradation. The major routes of drug degradation have been thoroughly reviewed.[45,46]

Although determination of a complete pH-degradation rate profile is desired, it may not always be practical due to limitations of drug supply and time. Also, insufficient solubility in purely aqueous systems may limit determination of pH-degradation rate profiles. Organic cosolvents may be used to increase solubility; however, extrapolation to aqueous conditions must be done with caution. Stability of the drug in a suspended form in the desired buffer can be tested in lieu of solution stability. The stress test results must however, be interpreted in relation to the solubility in the suspension medium. The test may provide an empirical indication of pH stability in the presence of excess water. Satisfactory stability in the GI pH range (1 to 7.5) is important for oral absorption. While there are examples of

successful solid oral dosage forms of drug that are highly unstable at GI pH (didonasine, esomaprazole magnesium[47]), excessive degradation in the GI tract limits oral absorption and may require heroic efforts to address the problem via formulation. Although no hard and fast rule exists on acceptable GI stability, Balbach and Korn[37] considered <2 to 5% degradation under simulated *in vivo* conditions (37°C, pH 1.2 to 8, fed and fasted conditions) as acceptable. Higher degradation may require additional investigation. The effect of the GI enzymes on drug stability should also be evaluated.

3.4.2 Preformulation Activities: Dependent on Solid Form

A new drug substance may exist in a multitude of crystalline and salt forms with different physical properties such as shape, melting point, and solubility that can profoundly impact the manufacturing and performance of its dosage form.

3.4.2.1 Solubility

Solubility is highly influenced by the solid-state form (e.g., crystalline or amorphous) of the drug. Rigorous solubility studies using the final solid form (i.e., salt form or crystal form) as a function of temperature (i.e., 25 and 37°C) and pH (range 1 to 7.5) are conducted during preformulation. Solubility in nonaqueous solvents is also screened. Solubility in simulated gastrointestinal fluids is also important.

For accurate determination of solubility:

- Attainment of equilibrium must be ensured by analyzing solution concentration at multiple time points until the concentration does not change considerably (i.e., <5% change in concentration).
- The pH of the saturated solution must be measured.
- The solid phase in equilibrium with the saturated solution must be analyzed by techniques such as hot stage microscopy, differential scanning calorimetry, or powder x-ray diffraction, to verify if the starting material has undergone a phase transformation.

3.4.2.2 Salt-Form Selection

The selection of an optimal chemical and physical form is an integral part of the development of an NCE. If an NCE is neutral or if its pK_a value(s) is not conducive to salt formation, it has to be developed in the neutral form (unless a prodrug is synthesized) and the only form selection involves the selection of its physical (crystal) form. However, if it exists as a free acid or a free base, then the "form" selection involves the selection of both chemical and physical forms. A decision must be made whether a salt or its free acid or base form should be developed. As will be described in Section 3.5, a salt form may lead to a higher dissolution rate and higher bioavailability for a poorly water-soluble drug. For a drug with adequate aqueous solubility, a salt form may not be necessary, unless, of course, a salt provides an advantage with respect to its physical form. In the pharmaceutical industry, salt selection is usually performed by a multidisciplinary team comprising representatives from the drug discovery, chemical development, pharmaceutical development, ADME, and drug safety departments. Serajuddin and Pudipeddi[27] reported that the following questions need to be satisfactorily addressed by the team in the selection of an optimal salt form for a compound: "Is the acid or base form preferred because of biopharmaceutical considerations? Is the salt form more suitable? Is the preparation of stable salt forms feasible? Among

various potential salt forms of a particular drug candidate, which has the most desirable physicochemical and biopharmaceutical properties?" With respect to physical properties, questions involve whether the compound exists in crystalline or amorphous form, and, if crystalline, whether it exhibits polymorphism.

At the outset of any salt selection program, it is important to determine whether the salt formation is feasible for the particular compound and, if yes, what counterions are to be used? Although it is generally agreed that a successful salt formation requires that the pK_a of a conjugate acid be less than the pK_a of the conjugate base to ensure sufficient proton transfer from the acidic to the basic species, the salt formation still remains a "trial and error" endeavor. Hundreds of individual experiments for salt formation of a particular compound are not uncommon. Because of the availability of HTS screening techniques in recent years there is no pressure to limit the number of such experiments. Serajuddin and Pudipeddi[27] reported that the number of feasibility experiments can be greatly reduced by studying the solubility vs. pH relationship of the drug and identifying the pH_{max} (the pH of maximum solubility). The nature of the pH–solubility profile and the position of pH_{max} depends on pK_a, intrinsic solubility (solubility of unionized species), and the solubility of any salt (K_{sp}) formed. For a basic drug, the pH must be decreased below the pH_{max} by using the counterion for a salt to be formed, and, for an acidic drug, the pH must be higher than the pH_{max}. Any counterion that is not capable of changing the pH in this manner may be removed from consideration. While salts may be formed from organic solvents by counterions that are not capable of changing the aqueous pH in this manner, such salts may readily dissociate in an aqueous environment. When the synthesis of multiple salts for a compound is feasible, the number may be narrowed down and the optimal salt may ultimately be selected by characterizing physicochemical properties of solids according to a multitier approach proposed by Morris et al.[48]

3.4.2.3 Polymorphism

Polymorphism is defined as the ability of a substance to exist as two or more crystalline phases that have different arrangements or conformations of the molecules in the crystal lattice. Many drug substances exhibit polymorphism. The definition of polymorphism according to the International Conference on Harmonization (ICH) guideline Q6A[49] includes polymorphs, solvates, and amorphous forms. Amorphous solids lack long-range order and do not possess a distinguishable crystal lattice. Solvates are crystal forms containing stoichiometric or nonstoichiometric amounts of solvent in the crystal. When the solvent is water they are termed hydrates. A thorough screening of possible crystal forms is conducted during candidate lead selection or shortly thereafter.

Typical methods for generation of polymorphs include sublimation, crystallization from different solvents, vapor diffusion, thermal treatment, melt crystallization, and rapid precipitation. High-throughput screening methods have been reported for polymorph screening.[50]

Methods for characterization of polymorphs include crystallographic techniques (single crystal and powder x-ray diffraction), microscopic characterization of morphology, thermal characterization (DSC/TGA), solution calorimetry, solid-state spectroscopic methods (IR, Raman, NMR), and solubility and intrinsic dissolution rate methods. Of these, the relative solubility or intrinsic dissolution rate is directly related to the free energy difference, and, hence the relative stability of polymorphs. Thermal data can also be used to assess relative stability of polymorphs. The form with the lowest solubility and, hence, free energy is the most stable form at a given temperature. Other forms would eventually transform to the stable form. The kinetics of crystal nucleation and growth determines the crystal form obtained during crystallization. Sometimes metastable forms are more readily crystallized than the most stable (and often desired) form. The kinetics of transformation of a metastable form to the stable form may be very slow and unpredictable. The

unexpected appearance of a less soluble polymorphic form of a marketed antiviral drug, ritonavir, caused serious problems with manufacture of its oral formulation.[51] The crystal form can have a profound influence on physicochemical properties of a drug substance. Melting point, solubility, stability, and mechanical properties may depend on the crystal form. A difference in solubility of crystal forms may manifest as a difference in bioavailability but the impact would depend on the dose, solubility of each form, and permeability. Formulation factors such as particle size and wettability often complicate the observed difference in the bioavailability of crystal forms. Polymorphism of pharmaceutical solids has been extensively studied.[52]

3.4.2.4 Solid-State Stability

The shelf-life of a product is often predicted by the stability of the drug substance. A thorough investigation of the drug substance stability is, therefore, necessary during preformulation following the identification of the final salt or crystal form. Typical stability testing during preformulation includes accelerated testing for 2 to 4 weeks at high temperatures (50 to 80°C) in dry or moist (75% RH) conditions. Photostability is also conducted as per ICH light sensitivity testing guidelines.[53] The criteria for an acceptable solid-state stability are compound-specific. Balbach and Korn[37] recommended a degradation of <3 to 5% at 60°C (dry) and <10 to 20% at 60°C (100% RH) in 2 weeks as acceptable. Physical stability (i.e., change in crystal form) must also be investigated under accelerated temperature and humidity conditions.

3.4.2.5 Drug-Excipient Interactions

The stability of the pure drug is often greater than when it is formulated. Excipients may facilitate moisture transfer in the drug product or initiate solid-state reactions at the points of contact and adversely impact the stability of the drug substance. For the same reason, it is not uncommon for the drug product stability to decrease when the drug concentration in the formulation is reduced.[54] Pharmaceutical excipients can interact chemically with the active pharmaceutical ingredient. Drug-excipient compatibility studies are conducted during preformulation to select the most appropriate excipients for formulation. The following classes of excipients for oral dosage forms are commonly employed: diluents or fillers, binders, disintegrants, glidants, colors, compression aids, lubricants, sweeteners, preservatives, suspending agents, coatings, flavors, and printing inks.

A typical drug-excipient compatibility study includes preparation of binary mixtures of the drug and excipients in glass vials. The ratio of the drug to each excipient must be comparable to that in the formulation. It is prudent to set up two sets of samples to bracket the drug concentration foreseen in the final products. Multicomponent mixtures of drug and excipients mimicking prototype formulations can also be made and tested under accelerated conditions. The mixtures are subjected to accelerated stress conditions for a period of 2 to 4 weeks. The samples are analyzed by HPLC for the percent drug remaining and any degradation products formed. Samples are typically stored at a reference condition (-20 or 5°C) and at elevated temperatures of 40 to 60°C under dry and moist conditions. Moist conditions are usually obtained by storing samples at 75 or 100% RH, or by adding 10 to 20% (w/w) water to the mixtures. Serajuddin et al.[55] described the principles and practice of drug-excipient compatibility studies for selection of solid dosage-form composition. Calorimetric methods for drug-excipient compatibility have also been described.[56]

Whether physical mixtures represent real drug-excipient(s) interactions in a capsule or tablet formulation has been debated.[57] The use of prototype formulations instead of physical mixtures has been suggested. However, drug substance availability and practicality may limit

such an approach. Although simplistic, a well-conducted drug-excipient compatibility study using physical mixtures can lead to a good understanding of drug stability and robust formulations.[55]

One must also consider that pharmaceutical excipients can possess acidity (or basicity). Although the chemical nature of excipients (i.e., mannitol, starch, sucrose) may itself be neutral, trace levels of acidic or basic residuals from the manufacturing process often impart acidity or basicity to excipients. Depending on the relative proportion of the drug and its acid/base nature, the microenvironmental pH of the drug product can be influenced by the excipient acidity or basicity. For example, addition of 2% magnesium stearate to a low drug content formulation can result in a microenvironmental pH above 8. An approximate but simple way to measure the microenvironmental pH of a drug-excipient mixture is to measure the pH of a slurry prepared with minimal amount of water. Excipients may also contain impurities that can accelerate drug degradation. For example, the following impurities may be present in the excipients listed with them: aldehydes and reducing sugars in lactose, peroxides and aldehydes in polyethylene glycol, heavy metals in talc, lignin and hemicellulose in microcrystalline cellulose, formaldehyde in starch, and alkaline residues such as magnesium oxide in stearate lubricants.[58] The physical state of an excipient (i.e., particle size,[59] hydration state,[60] and crystallinity[61]) can also affect drug stability. It is, therefore, important to review the excipient literature data, such as *Handbook of Pharmaceutical Excipients*,[62] prior to its use in a formulation.

In assessing drug-excipient compatibility, the same guidelines as in solid-state stability can be used. Typically degradation of less than 5% at high-temperature and humidity conditions (50 to 80°C, 75% RH) is considered acceptable, and higher degradation requires additional investigation under more moderate conditions.

3.4.2.6 Powder Properties of Drug Substance

For solid oral dosage forms, powder properties such as powder flow, density, and compactibility are important. For products with low drug concentration (e.g., <5 to 10%), the powder properties of the drug substance are usually less influential than excipients on the overall mechanical properties of the formulation. For products with high drug concentration (i.e., >50%, w/w), powder properties of the drug substance (or active pharmaceutical ingredient, API) may have significant influence on processability of the formulation. Since pharmaceutical scientists often manage to find engineering solutions to address poor powder properties of materials, there are no hard and fast rules for acceptable powder properties. However, consideration of powder properties from the early stages of development can result in more optimized and cost-saving processes. Powder properties of interest include:

- Particle morphology
- Particle size and particle size distribution, surface area
- True and relative densities
- Compaction properties
- Powder flow properties

Hancock et al.[63] reported a comprehensive review of a wide variety of pharmaceutical powders. Investigation of electrostatic properties of the drug substance or drug-excipient mixtures during preformulation has been recommended.[64] Such studies may be particularly relevant for dry powder inhalation systems.

Optimal compression or compaction properties of powders are critical for a robust solid dosage form. Although prediction of compaction properties of powders is not fully

possible, tableting indexes[65] and other material classification methods[66] have been proposed to assess compaction properties of powders. It is clear that no universal method exists to predict processing performance of powders. However, examination of powder properties during preformulation and comparing them with generally well-behaved materials from past experience can identify undesirable powder properties early on and even identify remedial methods such as modification of power properties through changes in crystallization techniques.

3.5 Biopharmaceutical Considerations in Dosage-Form Design

For systemic activity of a drug molecule, it must be absorbed and reach the bloodstream or the site of action if given by the oral, topical, nasal, inhalation, or other route of administration where a barrier between the site of administration and the site of action exists. Even when a drug is injected intravenously or intramuscularly, one must ensure that it is not precipitated at the site of administration and that it reaches the site of action. In the development of dosage forms for a particular drug, a formulator must, therefore, carefully consider various physicochemical, biopharmaceutical, and physiological factors that may influence absorption and transport of drugs. Appropriate formulation strategies must be undertaken to overcome the negative influences of any of these factors on the performance of dosage forms.

A large majority of pharmaceutical dosage forms are administered orally, and, in recent years, the drug solubility has become the most difficult challenge in the development of oral dosage forms. For example, in the 1970s and 1980s, when dissolution, bioavailability, and bioequivalence of drugs came under intense scrutiny and many of the related FDA guidelines were issued, a drug with solubility less than 20 μg/mL was practically unheard of. Presently, new drug candidates with intrinsic solubility less than 1 μg /mL are very common. In addition to solubility, physicochemical factors influencing oral absorption of drugs include dissolution rate, crystal form, particle size, surface area, ionization constant, partition coefficient, and so forth. Among the physiological factors, drug permeability through the GI membrane is of critical importance. Other physiological factors playing important roles in the performance of an oral dosage form are transit times in different regions of the GI tract, GI pH profile, and the presence of bile salts and other surfactants. Physiological differences such as the unfed vs. the fed state also need to be considered.

Since solubility and permeability are the two most important factors influencing oral absorption of drugs, the following biopharmaceutical classification system (BCS) for drug substances, based on the work by Amidon et al.,[67] has been recommended by the FDA:[68]

- Class I — Drug is highly soluble and highly permeable
- Class II — Drug is poorly soluble, but highly permeable
- Class III — Drug is highly soluble, but poorly permeable
- Class IV — Drug is both poorly soluble and poorly permeable

For a BCS Class I compound, there are no rate-limiting steps in drug absorption, except gastric emptying, and, therefore, no special drug-delivery consideration may be necessary to make the compound bioavailable. On the other hand, for a BCS Class II compound, appropriate formulation strategy is necessary to overcome the effect of low solubility. For a BCS Class III compound, formulation steps may be taken to enhance drug permeability through the GI membrane, although the options could be

very limited. Often, the continued development of a BCS Class III compound depends on whether its low bioavailability from a dosage form because of poor permeability is clinically acceptable or not. A BCS Class IV compound presents the most challenging problem for oral delivery. Here, the formulation strategy is often related to enhancing the dissolution rate to deliver maximum drug concentration to the absorption site. As for a Class III compound, formulation options to enhance drug permeability are often limited. Because of the importance of BCS in the development of dosage-form design strategy, an early classification of new drug candidates is essential to identify their formulation hurdles.

Some of the physicochemical and physiological factors involved in the design of oral dosage forms are discussed below in more detail.

3.5.1 Physicochemical Factors

3.5.1.1 Solubility

Low or poor aqueous solubility is a relative term. Therefore, the solubility of a compound must be considered together with its dose and permeability. A simple approach to assess oral absorption with a drug substance could be the calculation of its maximum absorbable dose (MAD):[69]

$$\mathrm{MAD} = S \times K_a \times \mathrm{SIWV} \times \mathrm{SITT} \tag{3.5}$$

where S is solubility (mg/mL) at pH 6.5, K_a the transintestinal absorption rate constant (per min) based on rat intestinal perfusion experiment, SIWV the small intestinal water volume (250 mL), and SITT the small intestinal transit time (4 h). One limitation of the MAD calculation is that only the aqueous solubility in pH 6.5 buffer is taken into consideration. There are many reports in the literature where the "*in vivo* solubility" of drugs in the GI tract in the presence of bile salts, lecithin, lipid digestion products, etc., was found to be much higher than that in the buffer alone. Therefore, MAD may be considered to be a conservative guide to potential solubility-limited absorption issues and whether any special dosage forms need to be considered to overcome such issues. Advanced software tools are available to estimate oral absorption.[70]

For an acidic and basic drug, the solubility over the GI pH range varies depending on the intrinsic solubility (S_o) of the compound (i.e., solubility of unionized or nonprotonated species), pK_a, and the solubility of the salt form.[71,72]

3.5.1.2 Dissolution

Dissolution rate, or simply dissolution, refers to the rate at which a compound dissolves in a medium. The dissolution rate may be expressed by the Nernst–Brunner diffusion layer form of the Noyes–Whitney equation:

$$J = \frac{\mathrm{d}m}{A\,\mathrm{d}t} = \frac{D}{h}(c_s - c_b) \tag{3.6}$$

where J is the flux, defined as the amount of material dissolved in unit time per unit surface area (A) of the dissolving solid, D the diffusion coefficient (diffusivity) of the solute, c_s the saturation solubility that exists at the interface of the dissolving solid and the dissolution medium, and c_b the concentration of drug at a particular time in the

bulk dissolution medium. Under "sink" conditions ($c_b < 10\% c_s$), the above equation is reduced to

$$J \approx \frac{D}{h} c_s \tag{3.7}$$

From the above equation, the dissolution rate is proportional to solubility at the early stage of dissolution. However, there is an important distinction between dissolution and solubility; solubility implies that the process of dissolution has reached equilibrium and the solution is saturated.

The importance of dissolution on drug absorption and bioavailability may be described in relation to the concept of dissolution number, D_n, introduced by Amidon et al.[67] D_n may be defined as

$$D_n = \frac{t_{res}}{t_{diss}} \tag{3.8}$$

where t_{res} is the mean residence time of drug in the GI tract and t_{diss} the time required for a particle of the drug to dissolve.

It is evident from the above equation that the higher the D_n, the better the drug absorption, and a maximal drug absorption may be expected when $D_n > 1$, i.e., $t_{res} > t_{diss}$. However, for the most poorly water-soluble drugs, $D_n < 1$ and, as a result a dissolution rate-limited absorption is expected. To ensure bioavailability of poorly soluble drugs, formulation strategies must, therefore, involve increasing the dissolution rate such that the full dose is dissolved during the GI residence time, i.e., D_n becomes equal to or greater than 1.

Some of the current methods of increasing dissolution rates of drugs are particle size reduction, salt formation, and development of the optimized delivery systems, such as solid dispersion, soft gelatin encapsulation, etc.

The effect of particle size on drug absorption as a function of dose and drug solubility was analyzed by Johnson and Swindell,[73] where the absorption rate constant was assumed to be 0.001 min^{-1}. The results are shown in Figure 3.1; Figure 3.1(a) shows that for dissolution rate-limited absorption ($D_n < 1$), the fraction of drug absorbed at any particle size will decrease with the increase in dose. On the other hand, if the dose is kept constant, the fraction of drug absorbed will increase with an increase in drug solubility, and the particle size becomes practically irrelevant for drugs with a solubility of 1 mg/mL at a dose of 1 mg (Figure 3.1(b)).

Salt formation increases the dissolution rate by modifying pH and increasing the drug solubility in the diffusion layer at the surface of the dissolving solid.[74] Depending on the pH in the GI fluid, a drug may precipitate out in its respective free acid or base form; however, if redissolution of the precipitated form is relatively rapid,[75] faster drug absorption is expected from the salt form. Precipitation, in the presence of food, may confound this expectation. Although salt formation increases drug dissolution rate in most cases, exceptions exist. Under certain situations, a salt may convert into its free acid or base form during dissolution directly on the surface of the dissolving solid, thus coating the drug surface and preventing further dissolution at a higher rate. In such a case, the salt formation may not provide the desired advantage and a free acid or base form may be preferred.[76]

A DDS such as a solid dispersion increases dissolution rate by dispersing the drug either molecularly or in an amorphous state in a water-soluble carrier.[9] When given as a solution in a pharmaceutically acceptable organic- or lipid-based solvent that is either

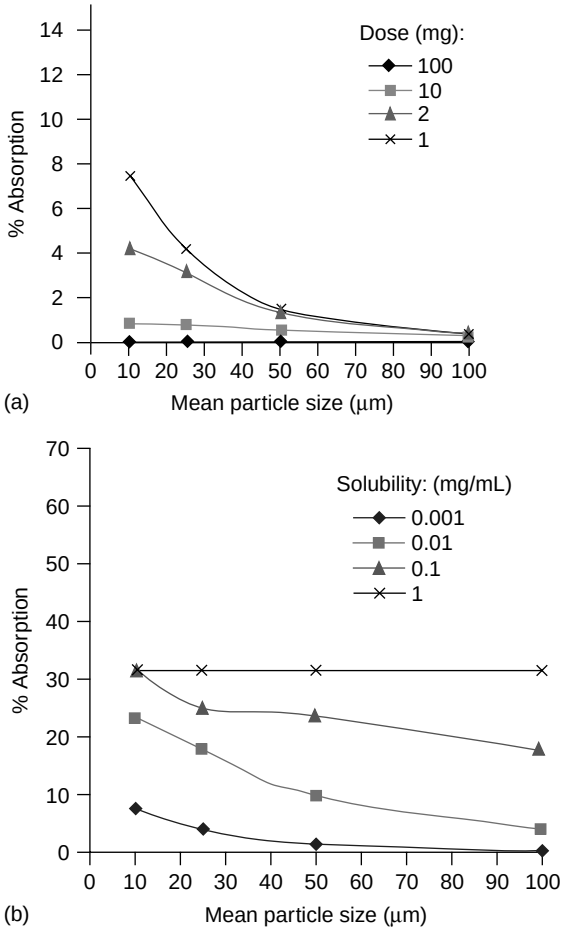


FIGURE 3.1
Computed percent of dose absorbed at 6 h vs. mean particle size, with an absorption rate constant of 0.001 min^{-1}: (a) at doses from 1 to 100 mg, with a solubility of 0.001 mg/mL; and (b) at solubility from 0.001 to 1.0 mg/mL, with a dose of 1 mg.

encapsulated in a soft gelatin capsule or packaged in a bottle, the drug immediately saturates the GI fluid after dosing and the excess drug may precipitate in a finely divided state that redissolves relatively rapidly.[77]

3.5.2 Physiological Factors

In any formulation development, the GI physiology that can influence dosage-form performance must be kept in mind. The GI contents play important roles, since even in the fasted state in humans, the *in vivo* dissolution medium is a complex and highly variable milieu consisting of various bile salts, electrolytes, proteins, cholesterol, and other lipids. The GI pH is another factor that plays a critical role in the performance of the dosage form. The pH gradient in humans begins with a pH of 1 to 2 in the stomach, followed by a broader pH range of 5 to 8 in the small intestine, with the intermediate range of pH values of around 5 to 6 being found in the duodenum. Colonic absorption in the last segment of the GI tract occurs in an environment with a pH of 7 to 8. The average pH values significantly differ between the fed and the fasted state. Finally, gastric emptying time and intestinal transit time are very important for drug absorption. The majority of the liquid gets emptied from the stomach within 1 h of administration.[78] Food and other solid materials, on the other hand, takes 2 to 3 h for half of the content to be emptied. The general range of the small intestinal transit time does not differ greatly and usually ranges from 3 to 4 h.[79]

3.5.2.1 Assessment of In Vivo Performance

Different formulation principles, dosage forms, and DDSs are commonly evaluated in animal models, and attempts are made to predict human absorption on the basis of such studies.[80] Human studies are also conducted in some cases to confirm predictions from animal models. Chiou et al.[81,82] demonstrated that there is a highly significant correlation of absorption ($r^2 = 0.97$) between humans and rats with a slope near unity. In comparison, the correlation of absorption between dog and human was poor ($r^2 = 0.512$) as compared to that between rat and human ($r^2 = 0.97$). Therefore, although dog has been commonly employed as an animal model for studying oral absorption in drug discovery and development, one may need to exercise caution in the interpretation of data obtained.

3.5.2.2 In Vitro–In Vivo *Correlation*

The various biopharmaceutical considerations discussed above can help in correlating *in vitro* properties of dosage forms with their *in vivo* performance.[83] *In vitro–in vivo* correlation (IV–IVC), when successfully developed, can serve as a surrogate of *in vivo* tests in evaluating prototype formulations or DDSs during development and reduce the number of animal experiments. This may also help in obtaining biowaivers from regulatory agencies, when applicable, and especially for BCS Class II compounds mentioned above. It can also support and validate the use of dissolution methods and specifications during drug development and commercialization.

The IV–IVC is generally established by comparing *in vitro* dissolution of drug with certain *in vivo* PK parameters. There are certain FDA guidelines for this purpose, where the correlations are categorized as Level A, Level B, Level C, and multiple Level C correlations.[84]

Level A correlation is generally linear and represents the correlation of *in vitro* dissolution with the drug fraction absorbed, which is obtained from the deconvoluted *in vivo* plasma levels. The Level B correlation is more limited in nature, and it does not uniquely reflect the actual *in vivo* plasma level curve; here, the mean *in vitro* dissolution time is compared with the mean *in vivo* residence or dissolution time. A Level C correlation establishes a single-point relationship between a dissolution parameter and a pharmacokinetic parameter. While a multiple Level C correlation relates one or several pharmacokinetic parameters of interest to the amount of drug dissolved at several time points of the dissolution profile, a Level C correlation can be useful in the early stages of formulation development when pilot formulations are being selected. For the purpose of IV–IVC, it is essential that the dissolution medium is biorelevant such that *in vitro* dissolution testing is indicative of *in vivo* dissolution of the dosage form. Dressman[85] published extensively on the identification and selection of biorelevant dissolution media.

3.6 Clinical Formulation Development: Clinical Trial Materials

A pharmaceutical dosage form is a means to deliver the active ingredient in an efficacious, stable, and elegant form to meet a medical need. The requirements and specifications of the product will depend on the scope and extent of its intended use in the various clinical phases and then through commercialization. Once the decision is made to move a drug candidate into clinical studies, the interaction of formulation scientists begins to shift from preclinical groups (medicinal chemistry, ADME) to those from development (analytical, process chemistry, clinical), marketing, and operations groups.

The formulation, manufacture, packaging, and labeling of CTMs precede the individual phases of clinical development.

Clinical development of NCEs is broadly divided into Phases I, II, III, and then IV (post-marketing) studies.

Phase I studies evaluate the pharmacokinetics and safety of the drug in a small number (tens) of healthy volunteers. Phase I studies are sometimes conducted in a small patient population (Proof of Concept studies) with a specific objective such as the validation of the relevance of preclinical models in man. The purpose of these studies may be the rapid elimination of potential failures from the pipeline, definition of biological markers for efficacy or toxicity, or demonstration of early evidence of efficacy. These studies have a potential go/no-go decision criteria such as safety, tolerability, bioavailability/PK, pharmacodynamics, and efficacy. Dosage forms used in Phase I or Proof of Concept studies must be developed with the objectives of the clinical study in mind.

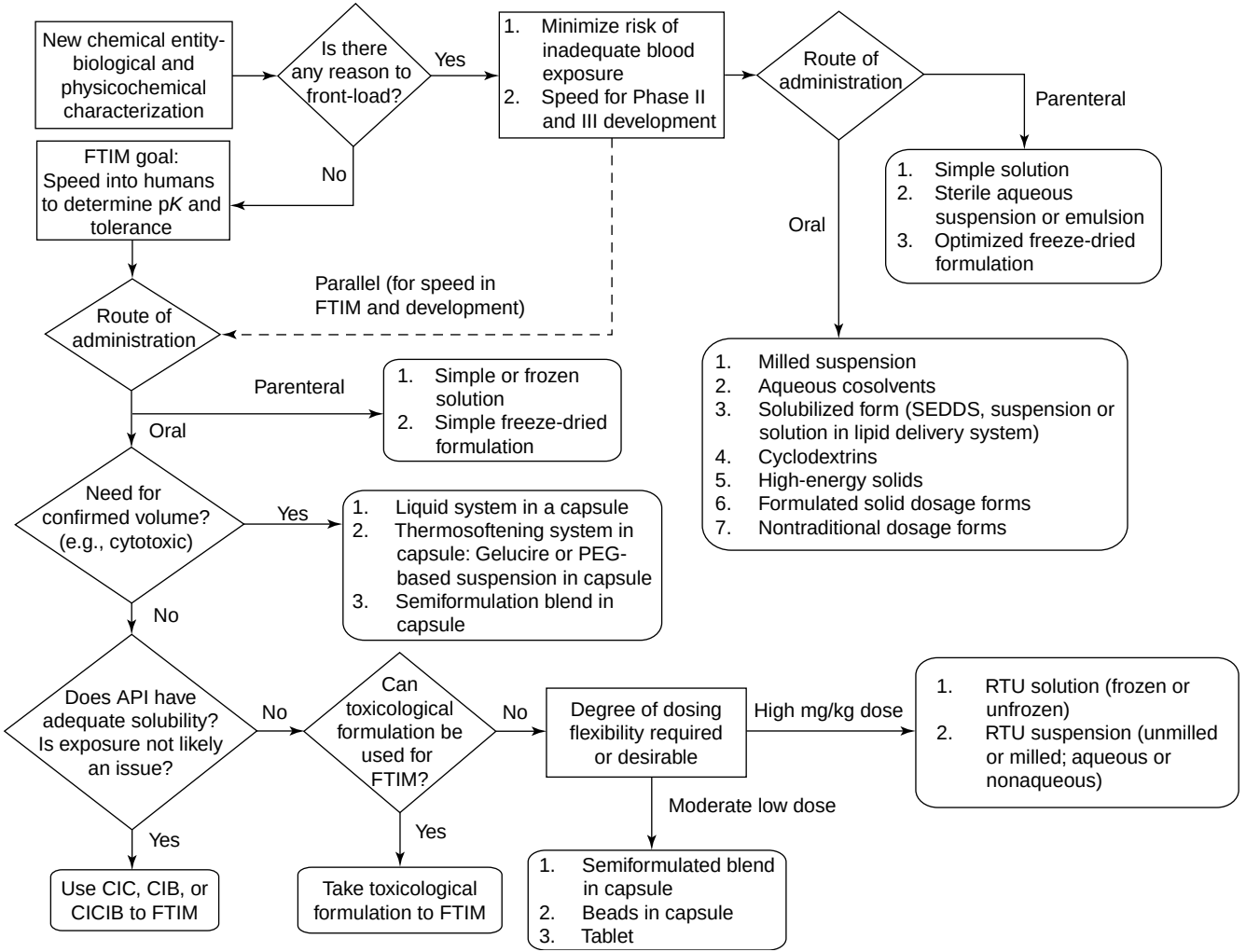
Phase II studies encompass a detailed assessment of the compound's safety and efficacy in a larger patient population (a few-to-several hundreds of patients). It is important that any formulation selected for these studies must be based on sound biopharmaceutical and pharmaceutical technology principles. Phase III clinical studies, also referred to as pivotal studies, involve several thousands of patients in multiple clinical centers, which are often in multiple countries. The aim of these studies is to demonstrate long-term efficacy and safety of the drug. Since these studies are vital in the approval of the drug, the dosage form plays a very critical role.

3.6.1 Phase I Clinical Trial Material

A decision tree approach for reducing the time to develop and manufacture formulations for the first oral dose in humans has been described by Hariharan et al.[86] and is reproduced in Scheme 3.1. The report summarized numerous approaches to the development and manufacture of Phase I formulations. Additional examples of rapid extemporaneous solution or suspension formulations for Phase I studies have been reported.[87,88]

In deciding the appropriate approach for early clinical studies, it is important to consider the biopharmaceutical properties of the drug substance and the goals of the clinical study. Practical considerations, such as the actual supply of the bulk drug and the time frame allotted for development, enter into the picture. The advantage of using extemporaneous formulations is the short development timelines (a few months) and the minimal drug substance requirements (a few hundred grams depending on the dose). Additional benefits include high-dose flexibility, minimal compatibility or formulation development, and minimal analytical work. The disadvantages include possible unpleasant taste, patient compliance issues, and dosing inconvenience for multiple-dose studies. For poorly soluble compounds, use of a nonaqueous solution may result in high systemic exposure that may be difficult to reproduce later on with conventional formulations. The "drug-in-capsule" approach, where the neat drug substance is encapsulated in hard gelatin capsules, has similar time, material, and resource advantages but is limited to compounds that exhibit rapid dissolution.

An intermediate approach for Phase I CTM is a "formulated capsule" approach, where the preformulation and compatibility data are optimally utilized and a dosage form that has scale-up potential is developed. The pharmaceutical composition is chosen on the basis of excipient compatibility and preliminary *in vitro* dissolution data. Extensive design of experiments is not conducted at this stage. Processing conditions are derived from past experience or simple reasoning. Several hundred grams of the drug substance are typically necessary to develop a formulated capsule product. However, the advantage



SCHEME 3.1
A decision tree approach for the development of first human dose. FTIM = first in man; CIC = chemical in capsule; CIB = chemical in bottle; CICIB = chemical in capsule in bottle; RTU = ready to use. Reproduced from Hariharan, M. et al., *Pharm. Technol.*, 68, 2003.

of a well-formulated capsule product is that it may be processed on a low-to-medium speed automatic equipment to meet the demands of a larger clinical study, if needed. A formulated tablet cannot be ruled out for Phase I; there is anecdotal information that a few drug companies use this approach. However, tablet development activities require more material and resources. While formulation design activities are in progress, the development of analytical methods must be initiated to develop an assay to separate the drug from its excipients, a stability-indicating method, and other tests (content uniformity, dissolution) that will be required to release a batch for human use. Limited drug substance and drug product stability testing are required to support a shelf-life (typically a few months, preferably at room temperature or under refrigeration) sufficient to conduct the Phase I study.

3.6.2 Phase II Clinical Trial Material

Extemporaneous formulations such as "powder-in-bottle" are unlikely to meet the demands of Phase II. Formulated capsules or tablets are typically necessary. The same dosage form used in Phase I studies may be continued for Phase II studies, if all the required dosage strengths can be supported and medium-to-large-scale (100,000 or more units) manufacturing of the dosage form is feasible. Alternatively, development of a more robust formulation closer to the desired commercial form may be undertaken. The drug substance requirements for such a Phase II dosage form may be in the range of the tens of kilograms, depending on the dose. The chemical development of the drug substance must also be well advanced to produce large quantities of the drug substance with minimal change in the impurity profile that is used in the toxicology program. The design of the clinical study may also influence the type of the dosage form. For example, in a double-blind study, the dosage-form presentation must be designed in a way to mask the difference in appearance between various dose strengths. It is often necessary to prepare placebos or active controls that consist of a marketed product disguised (masked) to look like the CTM. This is not a trivial exercise as consideration must be given to insure any changes made to the commercial product in the blinding process do not alter its bioavailability or stability. A popular approach is to place the commercial product into a hard gelatin capsule that looks like the capsules used for the investigative drug. When this is not possible (due to size issues) then the "double-dummy" approach is required. Here, a look-alike placebo of the marketed product must be prepared, which is then used in various schemes with the investigative product to provide the required dose combinations. Preparation of look-alike products, however, can pose difficult problems when the tablet surfaces are embossed with logos.

For the development of robust formulation and process, critical formulation issues must be first identified from the preformulation work. Critical issues may include solubility or dissolution rate for poorly soluble drugs, drug stability and stabilization, or processing difficulties due to poor powder properties of the drug substance (for high-dose formulations). In the initial stages of product development, *in vitro* dissolution in physiologically meaningful media must be utilized to guide the development of the prototype formulations.[68,89] Biopharmaceutical support in formulation development has been reviewed.[90]

If drug stability is identified as a potential issue, more thorough and careful drug-excipient compatibility studies may be necessary. On the basis of results of forced degradation and compatibility studies, a stabilization mechanism may be identified. Yoshioka and Stella[46] have thoroughly reviewed drug stability and stabilization. Stabilization strategies such as incorporation of antioxidants,[91] protection from moisture, or use of pH modifiers[92] may be considered. If a "functional" excipient such as an antioxidant or a pH modifier is

required, then tests to monitor its performance (i.e., consumption of the antioxidant during storage) must be developed.

A pragmatic but critical factor in Phase II formulation development is the identification of the dosage strengths by the clinical development scientists. Since Phase II studies cover a large dose range, the number of dosage strengths may be quite large. The selection of minimal number of dosage strengths that can still support the needs of the clinical studies is best handled by a team comprising pharmaceutical scientists, pharmacokineticists, and the clinical study and clinical supply coordinators. Individual strategies must be developed for both the high- and low-dose strengths. The limitations on the high-dose strength may include dissolution, processability, and size of the unit. Whereas the low-dose strengths may face mass or content uniformity or stability issues. Once the critical formulation issues are identified, a systematic approach is necessary to identify the optimal composition and process. Statistical design of experiments provides a systematic basis for screening of the prototype formulations or variants. It is important to correctly identify the desired responses to be optimized (i.e., dissolution, chemical stability, processing of a unit operation). It is, however, not uncommon, due to lack of the sufficient drug substance, to use empirical approaches to the formulation development. Systematic trial and error[93] experiments can also result in the identification of optimum conditions in a semistatistical way. Each of the development steps such as formulation composition, identification, and processing unit operations can be the subject of statistical optimization. However, time and resource limitations often preclude using such multiple optimizations. Extensive statistical designs to optimize processing conditions are often deferred to at a later stage. Use of expert systems[94] to build consistency in design and test criteria has been described.

In developing Phase II formulations with scale-up potential, manufacturing equipment that operates on the same principle as the production equipment should be used as much as possible. A difference in the operating principle such as low shear vs. high shear granulation may yield different product characteristics.[95] About three-to-four variants are identified from the screening studies for stress stability testing. During Phase I and II clinical studies the duration of the stability study will depend on the extent of the clinical studies. One-to-two-year stability studies under ICH[96] recommended storage conditions are desired with both accelerated and real-time storage conditions. Stability studies monitor changes in drug product performance characteristics including potency, generation of degradation products, product appearance, and dissolution. Stability data from these studies are utilized to develop the final product and to select the most appropriate packaging conditions.

Hard gelatin capsules and tablets are among the most common dosage forms for Phase II clinical studies. Capsules are considered to be more flexible and forgiving than tablets for early clinical studies. Capsules may also offer an easy means of blinding as noted previously for the use of active controls. The filling capacity of capsules is limited (usually $<$400 mg fill weight of powders for a size 0 capsule) and large doses may require capsules that are not easy to swallow.

Production of tablets at medium-to-large scale requires more stringent control of powder properties due to the high-speed compression step. Processing of tablets and the physics of tablet compaction have been the subject of extensive investigation and voluminous literature exists on the topic.

Overall, the development of a robust formulation with scale-up potential for Phase II studies involves integration of physicochemical, biopharmaceutical, and technical considerations. Whether a rudimentary formulated capsule or a more robust formulation closer to the commercial form will be used in Phase II studies will depend on the company policy, material cost, the complexity of clinical design, and the development strategy.

3.6.3 Phase III Clinical Trial Material

The purpose of Phase III clinical studies is to demonstrate the long-term safety and efficacy of the drug product. To insure the same safety and efficacy in the commercial product, it is necessary that the Phase III CTM should be close to or identical to the final commercial form when possible. A brief overview of the regulatory requirements for the commercial dosage form may help contribute to the understanding of the strategies that can be used for Phase III CTM.[96] According to the ICH guidelines, data from stability studies should be provided on at least three primary batches of the drug product. The primary batches should be of the same formulation and packaged in the same container closure system as proposed for marketing. Two of the primary batches should be of pilot scale or larger (i.e., the greater of 1/10th of the production scale or 100,000 units) and the third batch may be of smaller (lab) scale. Production scale batches can replace pilot batches. The manufacturing processes for the product must simulate those to be applied to the production batches and should provide a product of the same quality. The stability study should typically cover a minimum period of 12 months at the long-term storage condition (i.e., 25°C/60% RH). As a result of these requirements, the formulation, manufacturing process, and packaging of the drug product must be finalized at about 1.5 to 2 or more years prior to the filing of a regulatory application for an NCE.

If the company policy is to use at least pilot-scale batches for Phase III clinical studies, the formulation and manufacturing process must be finalized prior to the initiation of the pivotal clinical studies. This activity may require large quantities of the drug substance (i.e., hundreds of kilograms depending on the dose and dosage strengths). Additionally, the synthesis of the drug substance must be established.[96] The development of Phase III CTM must occur in parallel to Phase II clinical studies. The advantage of this approach is that a robust and well-developed formulation is utilized in Phase III clinical studies and scale-up risk is reduced. Scale-up and process validation can occur subsequently as per regulatory guidelines.[97] The concepts of design for manufacturability and process analytics technologies[98] are best incorporated into the initial design phases to promote the generation of robust, cost-efficient processes. Alternatively, the dosage form may be scaled-up to production scale at the launch site and the final commercial dosage form may be utilized for Phase III studies. This approach requires even more of the drug substance and a fully validated drug substance synthesis. Depending on the project strategy and the availability of resources, this approach may eliminate the risk of scale-up-related delays in launch. A detailed description of product scale-up and validation are beyond the scope of this chapter, but extensive literature is available on this topic.[41]

If the company resources, project timelines, or the complexity of the pivotal study design do not allow for the development of a pilot or the production scale formulation, the Phase III CTM may be significantly different from the final commercial form. The dosage form used for Phase III must, however, be robust enough to provide an uninterrupted drug supply for clinical studies. The Phase I or II CTM with the appropriate modifications (i.e., dosage strengths, blinding, manufacturability) may be used to meet the specific goals of Phase III. When significant composition or process changes are made to the CTM, a biopharmaceutical assessment must be made on the potential impact on systemic exposure. Clinical studies to establish bioequivalence between the previously used CTM and the new CTM or with the intended commercial dosage form may be required. The systemic exposure from early clinical formulations serves as a benchmark for later formulation changes. Demonstration of the bioequivalence of the Phase III CTM with the commercial dosage form, if different, may be necessary before regulatory approval and also in the case of certain postapproval changes.[99] In a bioequivalence

study a test product (such as an intended commercial dosage form) is compared with a reference formulation (i.e., a Phase III CTM material or an innovator's product in case of pharmaceutical equivalents) according to the guidelines established by the regulatory authorities. In the case of formulation changes for a CTM, the systemic exposure is compared to the earlier benchmark formulations, and substitution by the new product may be made on the basis of the product's efficacy and safety. Two products may be assessed to be equivalent even if the bioequivalence criteria are not strictly met, provided the new product does not compromise the safety and efficacy established by the previous product.[100] It may be possible to obtain a waiver for bioavailability or bioequivalence of formulations, depending on the BCS classification and the availability of established IV–IVC as discussed in the previous sections.

3.7 Nonoral Routes of Administration

As noted earlier, this chapter is focused on the design and evaluation of oral dosage forms. A few thoughts are provided to introduce some of the different biopharmaceutical issues that confront the design of some nonoral dosage forms.

3.7.1 Parenteral Systems

Many drugs are administered as parenterals for speed of action because the patient is unable to take oral medication or because the drug is a macromolecule such as a protein that is unable to be orally absorbed intact due to stability and permeability issues. The U.S. Pharmacopoeia defines parenteral articles as preparations intended for injection through the skin or other external boundary tissue, rather than through the alimentary canal. They include intravenous, intramuscular, or subcutaneous injections. Intravenous injections are classified as small volume (<100 mL per container) or large volume (>100 mL per container) injections. The majority of parenteral dosage forms are supplied as ready-to-use solutions or reconstituted into solutions prior to administration. Suspension formulations may also be used,[101] although their use is more limited to a subcutaneous (i.e., Novolin Penfill®; NOVO Nordisk) or intramuscular (i.e., Sandostatin LAR Depot®; Novartis) injection. Intravenous use of disperse systems is possible but limited (i.e., Doxil® Injection; Ortho Biotec).

The decision to develop a commercial parenteral dosage form must be made during drug discovery itself because the developability criteria are different than those of an oral dosage form. Additionally, extensive preformulation studies must be conducted to fully understand the pH-solubility and pH-stability properties of the drug substance: equilibrium solubility and long-term solution stability are important for parenteral dosage-form development.[102,103] The drugs generated by the "biopharmaceutical" industry are typically macromolecules such as proteins. Macromolecular drugs are generated via genetic engineering/fermentation techniques and purified in an aqueous solution. The material at the final processing step is measured into a package for ultimate delivery to the patient. If further processing such as lyophilization is required to maintain adequate shelf-life of the macromolecule, the composition (excipients) of the final product is carefully selected to maintain the physical and chemical stability of the bulk drug. This requires the coordination of the protein/processing specialist and the formulator to insure that the requisite stabilizers, pH control, and preservatives are added to the broth in the final processing steps.

Subcutaneous and intramuscular injections are administered in small volumes as a bolus (1 to 4 mL). Intravenous injections can be given as bolus (typically <5 mL) with larger volumes administered by infusion. Due to the small volume of bolus injections, high drug concentrations (up to 100 to 200 mg/mL) may be required for administration of large doses. As a result, solubility enhancement is a major consideration for parenteral dosage forms. Solubilization principles for parenteral dosage forms have been reviewed.[104] Common approaches to enhance solubility include pH adjustment, addition of a cosolvent, addition of a surfactant, complexation, or a combination of these approaches. Ideally, parenteral dosage forms should have a neutral pH. Injection volume and potential for pain[105] at the site of injection must be carefully evaluated in the selection of the pH of the formulation.

Judicious selection of excipients for parenteral dosage forms is critical due to their systemic administration.[106,107] Excipients of a parenteral dosage form may have a significant effect on product safety including injection site irritation or pain. Permissible excipients for parenteral dosage forms are far less than those for oral dosage forms.

Parenteral products must be sterile, particulate-free, and should be isotonic. An osmolarity of 280 to 290 mOsmol/L is desirable. Slightly hypertonic solutions are permissible but hypotonic solutions must be avoided to prevent hemolysis. It is essential that any product for injection be sterilized by a validated process starting with Phase 1 CTM. The method of sterilization (heat, filtration, high energy) of the product must be carefully considered as it can have a profound influence on the stability of the drug. This is also true for the choice of the packaging components as they too can influence stability by releasing materials (such as vulcanizing materials from rubber stoppers) capable of interacting with the drug.

Strategies for development of CTMs of parenteral dosage forms depend on their end use. For example, the requirements of an intravenous formulation intended only to determine the absolute bioavailability of a solid dosage form are more flexible than that being designed for commercial use. For example, if the stability of a solution formulation is limited, it may be supplied frozen to conduct a "one-time" bioavailability study. On the other hand, a parenteral formulation for long-term clinical use may require more extensive development to overcome such liabilities. If the stability of a solution formulation is not suitable for the desired storage condition (room temperature or refrigeration), a lyophilized formulation may be necessary. Principles and practice of lyophilization have been reviewed.[108] Bioavailability and bioequivalence considerations of parenteral dosage forms (nonintravenous) have been reviewed.[109]

3.7.2 Inhalation Systems

To treat diseases of the upper airways, drugs have been formulated as dry powders for inhalation, as solutions for nebulization, or in pressurized metered dose inhalers with the goal of delivering the drug topically in the bronchial region. These systems are inherently complex and can be inefficient as they tend to produce and deliver large particles that lodge in the back of the throat.[110] The FDA considers inhalation dosage forms as unique as their performance is markedly influenced by the formulation, the design of the packaging (container and valve) as well as patient-controlled factors. The factors such as rate and extent of the breath can influence the delivered dose uniformity and the particle size distribution. An additional challenge for the formulator is the limited number of excipients that have been demonstrated to be safe when delivered into the lung, which must be carefully sourced and controlled.[111]

The formulator's primary focus is on how to produce particles of the requisite size and how to maintain them as such prior to and during administration. For solid drugs, the particles are targeted to be produced in the range of 1 to 5 μm by milling,

controlled precipitation, or spray drying. When a drug is sufficiently soluble in water its solutions can be processed to create droplets in the appropriate particle size range by nebulizers (solutions must be sterile). Such fine particles are subject to agglomeration due to static charges or humidity and require exquisite control of the environment (atmospheric such as humidity, and local such as excipients) during production, packaging, and storage as changes in the PSD can profoundly influence the efficiency of delivery. These constraints can have a negative impact on the cost of goods. The pulmonary humidity may also result in changes to the PSD as the particles travel the length of the airways.[112] A recent approach mitigates these concerns by producing the particles on demand via vaporization of the drug followed by condensation into a controlled PSD as the patient inhales.[113] This approach is capable of producing nanometer- or micron-sized particles.

In addition to local lung delivery, pharmaceutical scientists have noted the extensive surface area of the lower, gas-transport region of the lung, the alveolar region, and have been seeking a means of reproducibly administering potent therapeutic agents into this region for rapid transport into the systemic circulation thus bypassing the GI tract and liver. The pharmacokinetic profile from this noninvasive method mimics that of IV injection.[114] The ability of a vaporized (smoked) drug to exert a rapid CNS effect was well known to our great ancestors and such rapid delivery is currently under consideration for treatment of acute and episodic conditions such as migraine or breakthrough pain. There is currently great interest to deliver macromolecules such as insulin in this manner employing the drug in solution or dry powders.[115]

3.8 Drug-Delivery Systems

From the perspective of absolute efficiency, most dosage forms are failures. Far too much drug is poured into the body so that a few molecules can reach the site of the desired pharmacologic response.[116] As already noted, the drug must pass many anatomical and physiologic (pH, enzymes) barriers on its journey to the target. Formulators often talk of a magic bullet that, in their fondest dreams, delivers the drug to the exact site needed and no more. Since the 1960s, when pharmaceutical scientists began to understand how the physicochemical properties of a drug influenced the ability of the drug to dissolve, remain stable in a changing pH milieu, be transported across the epithelium by passive and active transport, and how it would fair as it came in contact with numerous metabolizing systems (biopharmaceutical factors), great strides have been made in designing and improving the performance of oral dosage forms. With time, similar biopharmaceutical factors were revealed for other routes of delivery that enabled novel dosage forms to be prepared. Examples include transdermal patches to deliver the drug into the systemic circulation over days and weeks from a single application or polymeric microspheres capable of prolonged residence time in muscle tissue from where they could slowly release the needed peptide. New materials of construction (excipients, carriers) were required to create these novel systems, which added to the formulators' palette more than the ethanol (for elixirs) and starch (for tablets) used for decades. Though, we must be cautious when we add a component to our delivery system to nonspecifically overcome a barrier that has been evolved through millennia, as this can induce safety problems, such as been found with the use of dermal penetration enhancers for use in transdermal delivery patches. Therefore, the safety of any new excipient has to be demonstrated via chronic toxicity studies before they can be seriously considered for use. However, as our understanding of

the anatomical microenvironments improves and new construction materials become available, so does our desire to further fashion a magic bullet.

A DDS, which is also referred to as "novel" DDS, is a relative term. What we consider common dosage forms today were novel a few decades ago; one notable example of novel formulations, as we know them today, was the development of the tiny, little time-pills termed Spansules® by a group at SK&F.[117] These provided prolonged release of drugs, which simplified dosing and attenuated adverse reactions. The use of pharmacokinetic data to guide the design of the optimized delivery system was at that time a breakthrough.

Given the complexity in the design and subsequent development of novel DDS, it is impossible for each Big Pharma company to internally master many of the technological breakthroughs. This has led to the formation of a new industry — DDSs and tools — with an estimated several hundred participating companies. Like their big brothers, these companies also function with multidisciplinary teams with thorough knowledge of the anatomical and molecular barriers that must be overcome and the ability to design strategies to surmount them. While "traditional" drug-delivery companies employed new techniques to produce dosage forms (e.g., osmotic pumps for controlled release, nanocrystal milling), many of the newer companies are working at the molecular level to create drugs with better transport properties.

Just as a new molecular entity must pass hurdles to demonstrate its safety and efficacy, so too must new delivery systems meet these requirements. Furthermore, the cost of goods of the new technology must be appropriate to the benefit provided. As noted earlier, there is a long path from the first formulation to one robust enough to routinely prepare commercial quantities. These processing issues are even more important and really daunting for a novel DDS that has never been made in more than test tube quantity. This was the situation confronting the liposome companies founded in the early 1980s to commercialize this carrier technology. The issues of the "three-esses" prevailed — stability, sterility, and scale-up — and had to be overcome to make the technology worthy of consideration as a new dosage form. It took a decade and tens of millions of dollars spent by several companies to accomplish the task — performed under cGMP.[118] The situation is even more (financially and emotionally) perilous when a new DDS is used to support the clinical evaluation of a new drug. Thus, the use of an optimized delivery system oftentimes has lagged the introduction of a new medicine.

According to an estimate in early 2000s, 12% of the pharmaceutical market comprises DDSs. This share will definitely increase in the future as it is evident from some of the considerations made above. Among the market share of various DDSs, oral dosage forms account for approximately 60%, inhalation products for 20%, transdermal about 10%, and injectables around 9%.[119]

Drug-delivery systems are essentially specialized dosage forms developed to overcome the limitations of conventional dosage forms, such as simple tablets, capsules, injectable solutions, etc. Some of the reasons behind the development of oral DDSs are listed below:

- Overcome drug developability issues
- Unfavorable lipophilicity
- Poor permeability
- Poor stability
- Overcome physiological hurdles
- Relatively short half-life
- First-pass effect
- GI residence time
- GI pH
- Drug stability in GI fluid
- Effect of food and other GI contents

- Enhance absorption rate and bioavailability
- Increase solubility
- Increase dissolution rate
- Increase patient convenience
- Once-a-day (compared to multiple-time dosing)
- Geriatric formulation
- Pediatric formulation
- Intellectual property
- New formulation principles
- Product and process improvement

Many different technologies to address these drug-delivery issues are either available or are emerging through the efforts of Big Pharma or various smaller drug-delivery companies. For certain specific applications, multiple (commodity) technologies may be available. For example, matrix tablets, coated tablets, osmotic release systems, single unit vs. multiple units, etc., are available for prolonged-release dosage forms. Of course, each of these systems has its own advantages and disadvantages.

The need for DDSs is possibly the greatest in the case of biotechnologically derived products that cannot be orally absorbed, such as peptides, proteins, and oligonucleotides.

Some of the major formulation issues with such products are:

- Invasive dosage forms (injection)
- Short half-life
- Poor stability of drugs or dosage forms
- Lack of drug targeting

Many different approaches are being applied to resolve these issues. Some of the techniques include biodegradable microspheres, PEGylation, liposomal delivery, electroporation, prodrug and conjugate formation, dry powder inhalers, supercritical fluid-based nanoparticles, viral and nonviral vectors, and so forth. Some of the successes in this area include first microsphere sustained release formulation of a peptide LHRH (Lupron Depot®; TAP/Takeda) in 1989; first PEGylated sustained release formulation of a protein adenosine deaminase (Adagen®; Enzon) in 1990; first microsphere sustained release formulation of a recombinant protein, human growth hormone (Nutropin Depot®; Genentech) in 2000; and first PEGylated sustained release formulation of a recombinant protein, interferon α-2b (PEG-Intron®; Schering-Plough) in 2000.

It is impossible to describe the breadth of drug-delivery activities in one section of a chapter. The general approach is to focus on a particular "anatomical niche," i.e., use of the lung as a portal for systemic delivery of drugs and to develop technology to overcome the barriers as they are unveiled. Physiologic barriers can be overcome with knowledge of their molecular basis: use of inhibitors to block CYP metabolism of a drug, use of specific permeability enhancers to facilitate transport across a membrane system, blockage of efflux transporters, and use of a prodrug to optimize the partition coefficient and transport mechanisms. Another approach is to take the notion that a particular technology may have DDS applicability, e.g., nanoparticles, and then seek applications. Additionally, as "new" diseases are identified such as age-related macular disease, new noninvasive methods of delivery are required to reach the (retinal) targets at the back of the eye.[120]

The following are among the many novel delivery system approaches currently under development:

- *Tablets*: fast dissolving, lipid-carrier based, slow releasing, disintegrate without water, float in the stomach (gastric retentive), buccal

- *Dermal*: patches, iontophoresis
- *Solubility enhancement systems*: lipid-based systems, nanoparticles, surfactants, semisolid formulations
- Oral methods for peptides
- *Inhalation*: peptide delivery, vaporization
- *Site-specific delivery*: liposomes, drug monoclonal antibody
- *Implants*: biodegradable polymers
- Drug-eluting stents

A new trend in the delivery of medicines is to employ a device component. This may be an implantable pump for insulin, a metallic stent coated with a drug, or unit capable of rapidly vaporizing a discrete dose for inhalation. Such products are regulated by the FDA as "combination" products and may be reviewed by multiple Centers within the Agency, which may require additional levels of documentation to support the product design.

Although DDSs are bringing better therapy to patients, certain concerns with respect to their application for NCEs remain. At an early stage of drug development, the project needs are not usually clear, and proof of concept is usually more important than the optimal therapy. The application of DDS usually leads to longer development time and increased resources. The DDSs are usually more difficult to scale-up. Moreover, if the NCE is dropped from development, all the efforts and resources applied toward the development and application of the DDSs are wasted. For these reasons, decisions are sometimes made to develop initial dosage forms for NCEs utilizing the convenient approaches and to postpone DDS until the LCM phase. However, there is also another school of thought in this area. Given the financial pressures of Big Pharma to introduce blockbusters, and the multitude of leads generated by the HTS methods described above, there are those[121] who advocate the use of new technologies and partnerships to enable the use of otherwise delivery-impaired molecules. These may be considered "compound-enabling strategies rather than just product-enhancement programs."

In selecting one DDS over another, technical feasibility must be considered. One consideration that is often overlooked during the development of the drug-delivery opportunities is the dose. For example, an oral tablet may be able to deliver as much as 1 g of drug, while the maximum limit for an inhaled dose could be 1 mg/dose and a transdermal patch 5 to 10 mg/day. Therefore, an inhalation product or transdermal patch could not be the substitute due to the poor bioavailability or first-pass metabolism of a 500 mg tablet. Even the maximum dose in a tablet could be limited if a larger amount of excipients is needed for controlled release, taste masking, etc. The size limitation usually does not permit a larger than 300 mg dose in a capsule. For injectables, 1 to 2 mcg/day could be the maximum dose for a monthly depot system, while 50 to 100 mg might be delivered intramuscularly or up to 750 mg by intravenous bolus administration, and as much as 1 g or even more may be administered by infusion. Another important consideration is the cost. The cost of manufacturing of a DDS should not be so high that the product is not commercially feasible.

3.9 Product Life-Cycle Management

It is estimated that only 1 out of 10 drug molecules that are selected for development and undergoes various preclinical and clinical development activities ultimately reaches the market. Because of such an attrition rate, drug companies are often forced to conserve

resources during initial development and bring a product to the market that may not have optimal pharmaceutical and clinical attributes. This approach also leads to faster availability of new therapies to patients. After the initial launch, further development activities leading to superior products, known as the product LCM, continue. The LCM may lead to new dosage forms and delivery systems, new dosing regimen, new delivery routes, patient convenience, intellectual property, and so forth. An LCM program is considered successful only if it leads to better therapy and patient acceptance. There are numerous examples of successful LCM through the development of prolonged-release formulations. Development of nifedipine (Procardia® XL, Pfizer), diltiazem (Cardizem CD®, Aventis), and bupropion HCl (Wellbutrin SR® and XL®, GSK) prolonged-release products that not only provided more convenient and better therapy to the patients by reducing dosing frequency but at the same time greatly increased sales of the products are well-known examples. Even old compounds like morphine and oxycodone were turned into significant products by the development of more convenient prolonged-release formulations (MS Contin® and Oxycontin®, respectively; Purdue Pharma).

Issues with the bioavailability of drugs due to their poor solubility or reduced permeability were discussed earlier. Many of the future LCM opportunities may come through bioavailability enhancement. As discussed earlier, solid dispersion, microemulsion, soft gelatin capsule formation, solubilization, lipid-based DDSs, nanoparticle or nanocomposite formation, etc., are some of the common bioavailability approaches that can be utilized for LCM. Development of Lanoxicaps® by Burroughs-Wellcome in 1970s by encapsulating digoxin solutions in soft gelatin capsules is a classic example of LCM by bioavailability enhancement and better pharmacokinetic properties. The development of a microemulsion preconcentrate formulation by Novartis (Neoral®), where the variability in plasma and the effect of food were reduced greatly, is another well-known example.[122]

Life-cycle management through the development of fixed combination products, where two or more drugs are developed or copackaged into a single entity, is gaining increased popularity. The fixed combination products often provide synergistic effects, better therapy, patient compliance, patient convenience, increased manufacturing efficiency, and reduced manufacturing cost. However, a clear risk/benefit advantage is essential for the successful LCM by combination products; mere patient convenience may not be sufficient. Common justifications for the development of fixed combination products include improvement of activity such as synergistic or additive effect, improved tolerance by reduced dose of individual ingredients, broadening of activity spectrum, improvement of pharmacokinetic properties, and simplification of therapy for the patient.

The development of oral dosage forms that disintegrate or dissolve in the mouth is providing LCM opportunities for pediatric, geriatric, or bedridden patients who have difficulty in swallowing. They are also being used by active adult patients who may not have ready access to water for swallowing tablets or capsules.

3.10 Summary

This chapter describes the many approaches employed to identify and then develop new pharmaceuticals. This chapter has endeavored to outline the careful multidisciplinary tasks that must take place in a timely fashion to insure that the inherent power of each new drug candidate is allowed to be fulfilled. Since the first edition of this book was published, Big Pharma has become even bigger and the quest to regularly introduce blockbusters (products with sales in the $ billions) is even greater. This has resulted in the adoption of HTS tools in

discovery to increase the ability to produce and screen new molecules. As has been noted, the pressure befalling the product development scientists to make the right choices with regards to issues such as salt and polymorph selection at earlier time points has actually increased under this paradigm. Decisions regarding physical-chemical and physical-mechanical properties and their related impact on high-speed, efficient processing of the commercial dosage forms have to be made with little time and even less drug substance. This is remarkable in that a successful product will require a manufacturing process that is robust enough to consistently make millions of dosage units. The recent changes in the regulatory environment[123] have had a serious impact on the overall development and manufacturing process. The risk-based approach to Pharmaceutical c-GMP requires a complete understanding of pharmaceutical processes, which can only be achieved by the application of sound scientific principles throughout the discovery and development of a product. This chapter describes the underlying scientific principles of integrated product development, from lead selection to LCM.

References

1. Lemberger, A.P., Higuchi T., Busse, L.W., Swintowsky, J.V., and Wurster, D.E., Preparation of some steroids in microcrystalline form by rapid freeze-sublimation technique, *J. Am. Pharm. Assoc.*, 43, 338, 1954.
2. Nelson, E. and Schaldemose, I., Urinary excretion for evaluation of drug absorption I, *J. Am. Pharm. Assoc.,* 48, 489, 1959.
3. Levy, G. and Nelson, E., Pharmaceutical formulation and therapeutic efficacy, *JAMA,* 177, 689, 1961.
4. Gibaldi, M. and Kanig, J.L., Absorption of drugs through the oral mucosa, *J. Oral. Ther. Pharmacol.*, 31, 440, 1963.
5. Wagner, J.G., Biopharmaceutics: absorption aspects, *J. Pharm. Sci.*, 50, 359, 1961.
6. Liberman, S.V., Kraus, S.R., Murray, J., and Wood, J.H., Aspirin formulation and absorption rate I. Criteria for serum measurements with human panels, *J. Pharm. Sci.*, 53, 1486, 1964.
7. Kaplan, S.A., Biopharmaceutical considerations in drug formulation design and evaluation, *Drug Metab. Rev.*, 1, 15, 1972.
8. Yalkowsky, S.H. and Rubino, J.T., Solubilization by cosolvents Part 1, *J. Pharm. Sci.*, 74, 416, 1985.
9. Serajuddin, A.T., Solid dispersion of poorly water soluble drugs: early promises, subsequent problems, and recent breakthroughs, *J. Pharm. Sci.*, 88, 1058, 1999.
10. Lieberman, H.A., Lachman, L., and Schwartz, J.B., Eds., *Pharmaceutical Dosage Forms: Tablets*, Vols. I–III, Marcel Dekker, New York, 1990.
11. Gibson, M., *Pharmaceutical Preformulation and Formulation: A Practical Guide from Candidate Selection to Commercial Dosage Form*, CRC Press LLC, Leicestershire, 2004.
12. Venkatesh, S. and Lipper, R.A., Role of the development scientist in compound lead selection and optimization, *J. Pharm. Sci.*, 89, 145, 2000.
13. Drews, J., Drug discovery: a historical perspective, *Science*, 287, 1960, 2000.
14. Lipinski, C.A., Lombardo, F., Dominy, B.W., and Feeney, P.J., Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, *Adv. Drug Delivery Rev.*, 23, 3, 1997.
15. Prentis, R.A., Lis, Y., and Walker, S.R., Pharmaceutical innovation by the seven UK-owned pharmaceutical companies (1964–1985), *Br. J. Clin. Pharmacol.*, 25, 387, 1988.
16. Kerns, E.H. and Li, D., Pharmaceutical profiling in drug discovery, *Drug Disc. Today*, 8, 316, 2003.
17. Ho, N.F.H., Raub, T.J., Burton, P.S., Barsuhn C.L., Adson, A., Audus, K.L., and Borchardt, R.T., Quantitative approaches to delineate passive transport mechanisms in cell culture monolayers,

in *Transport Processes in Pharmaceutical Systems*, Amidon, G.L., Lee, P.I., and Topp, E.M., Eds., Marcel Dekker, New York, 2000, p. 219.
18. Clark, D.E., Computational methods for the prediction of ADME and toxicity, *Adv. Drug Delivery Rev.*, 54, 253, 2002.
19. Veber, D.F., Johnson, S.R., Cheng, H.Y., Smith, B.R., Ward, K.W., and Kopple, K.D., Molecular properties that influence oral absorption of drug candidates, *J. Med. Chem.*, 45, 2615, 2002.
20. Yalkowsky, S.H. and Valvani, S.C., Solubility and partitioning I: solubility of nonelectrolytes in water, *J. Pharm. Sci.*, 69, 912, 1980.
21. Klopman, G., Wang, S., and Balthasar, D.M., Estimation of aqueous solubility of organic molecules by the group contribution approach. Application to the study of biodegradation, *J. Chem. Inf. Comput. Sci.*, 32, 474, 1992.
22. Huuskonen, J., Salo, M., and Taskinen, J., Aqueous solubility prediction of drugs based on molecular topology and neural network modeling, *J. Chem. Inf. Comput. Sci.*, 38, 450, 1998.
23. Chen, X.-Q., Cho, S.J., Li, Y., and Venkatesh, S., Prediction of aqueous solubility of organic compounds using a quantitative structure–property relationship, *J. Pharm. Sci.*, 91, 1838, 2002.
24. McKenna, J.M., Halley, F., Souness, J.E., McLay, I.M., Pickett, S.D., Collis, A.J., Page, K., and Ahmed, I., An algorithm-directed two-component library synthesized via solid-phase methodology yielding potent and orally bioavailable p38 MAP kinase inhibitors, *J. Med. Chem.*, 45, 2173, 2002.
25. Kenseth, J.R. and Coldiron, S.J., High throughput characterization and quality control of small molecule combinatorial libraries, *Curr. Opin. Chem. Biol.*, 8, 418, 2004.
26. Faller, B., Improving solubility in lead optimization, *Am. Pharm. Rev.*, 7, 30, 2004.
27. Serajuddin, A.T.M. and Pudipeddi, M., Salt selection strategies, in *Handbook of Pharmaceutical Salts: Properties, Selection, and Use*, Stahl, P.H. and Wermuth, C.G., Eds., Wiley-VCH, Weinheim, 2002, p. 135.
28. Hilal, S.H., El-Shabrawy, Y., Carreira, L.A., Karickhoff, L.W., Toubar, S.S., and Rizk, M., Estimation of ionization pK_a of pharmaceutical compounds using the computer program SPARC, *Talanta*, 43, 607, 1996.
29. Albert, A. and Serjeant, E.P., *The Determination of Ionization Constants*, 3rd ed., Chapman & Hall, New York, 1984.
30. Donovan, S.F. and Pescatore, M.C., Method for measuring the logarithm of octanol-water partition coefficient by using short octadecyl-poly(vinyl-alcohol) high performance liquid chromatography columns, *J. Chromatogr. A*, 952, 47, 2002.
31. Comer, J. and Box, K., High throughput measurement of drug pK_a values for ADME screening, *JALA*, 8, 55, 2003.
32. Kansy, M., Senner, F., and Gubernator, K., Physicochemical high throughput screening, *J. Med. Chem.*, 41, 1007, 1998.
33. Hwang, K.K., Martin, N.E., and Jiang, L., Permeation prediction of M100240 using the PAMPA, *J. Pharm. Sci.*, 6, 315, 2003.
34. Artursson, P., Palm, K., and Luthman, K., Caco-2 monolayers in experimental and theoretical predictions of drug transport, *Adv. Drug Delivery Rev.*, 46, 27, 2001.
35. Stoner, C.L., Cleton, A., Johnson, K., Oh, D-M., Hallak, H., Brodfuehrer, J., Surendran, N., and Han, H.K., Integrated oral bioavailability projection using in vitro screening data as a selection tool in drug discovery, *Int. J. Pharm.*, 269, 241, 2004.
36. Ingels, F.M. and Augustijns, P.F., Biological, pharmaceutical, and analytical considerations with respect to the transport media used in the absorption screening system, Caco-2, *J. Pharm. Sci.*, 92, 1545, 2003.
37. Balbach, S. and Korn, C., Pharmaceutical evaluation of early development candidates—the 100 mg approach, *Int. J. Pharm.*, 275, 1, 2004.
38. Yu, H.S. and Adedoyin, A., ADME-Tox in drug discovery: integration of experimental and computational technologies, *Drug Discov. Today*, 8, 852, 2003.
39. Berry, I.R. and Nash, R., *Pharmaceutical Process Validation*, Marcel Dekker, New York, 1993.
40. Avdeef, A., Physicochemical profiling, *Curr. Top. Med. Chem.*, 1, 277, 2001.
41. Avdeef, A., Box, K.J., Comer, J.E.A., Gilges, M., Hibbert, C., Patterson, W., and Tam, K.Y., PH-metric log *P* 11. pK_a determination of water-insoluble drugs in organic solvent-water mixtures, *J. Pharm. Biomed. Anal.*, 20, 1999, 631.

42. Surakitbanharn, Y., McCandless, R., Krzyzaniak, J.F., Dannenfelser, R-M., and Yalkowsky, S.H., Self-association of dexverapamil in aqueous solution, *J. Pharm. Sci.*, 84, 720, 1995.
43. Hansch, C. and Leo, A., *Substituent Constants for Correlation Analysis in Chemistry and Biology*, Wiley Interscience, New York, 1979.
44. Reynolds, D.W., Fachine, K.L., Mullaney, J.F., Alsante, K.M., Hatajik, T.D., and Moto, G.M., Available guidance and best practices for conducting forced degradation studies, *Pharm. Technol.*, February, 48, 2002.
45. Carstensen, J.T. and Rhodes, C.T., Eds., *Drug Stability: Principles and Practices*, Marcel Dekker, New York, 2000.
46. Yoshioka, S. and Stella, V.J., *Stability of Drugs and Dosage Forms*, Kluwer Academic, New York, 2002.
47. Physician's Desk Reference 2004.
48. Morris, K.R., Fakes, M.G., Thakur, A.B., Newman, A.W., Singh, A.K., Venit, J.J., Spagnuolo, C.J., and Serajuddin, A.T.M., An integrated approach to the selection of optimal salt form for a new drug candidate, *Int. J. Pharm.*, 105, 209, 1994.
49. International Conference on Harmonization Q6A Guideline: Specifications for new drug substances and products: Chemical Substances, October 1999.
50. Morissette, S.L., Almarsson, O., Peterson, M.L., Remenar, J.F., Read, M.J., Lemmo, A.V., Ellis, S., Cima, M.J., and Gardner, C.R., High-throughput crystallization: polymorphs, salts, co-crystals and solvates of pharmaceutical solids, *Adv. Drug Delivery Rev.*, 56, 275, 2004.
51. Bauer, J., Spanton, S., Henry, R., Quick, J., Dziki, W., Porter, W., and Morris, J., Ritonavir: an extraordinary example of confirmation polymorphism, *Pharm. Res.*, 18, 859, 2001.
52. Brittain, H.G., *Polymorphism in Pharmaceutical Solids*, 1st ed., Marcel Dekker, New York, 1999.
53. ICH Guidance for Industry Q1B, Photostability testing of new drug substances and products, November 1996.
54. Badawy, S.I., Williams, R.C., and Gilbert, D.L., Chemical stability of an ester prodrug of a glycoprotein IIb/IIIa receptor antagonist in solid dosage forms, *J. Pharm. Sci.*, 88, 428, 1999.
55. Serajuddin, A.T.M., Thakur, A.B., Ghoshal, R.N., Fakes, M.G., and Varia, S.A., Selection of solid dosage form composition through drug-excipient compatibility testing, *J. Pharm. Sci.*, 88, 696, 1999.
56. Schmitt, E.A., Peck, K., Sun, Y., and Geoffroy, J-M., Rapid, practical and predictive excipient compatibility screening using isothermal microcalorimetry, *Thermochim. Acta*, 380, 175, 2001.
57. Monkhouse, D.C. and Maderich, A., Whither compatibility testing?, *Drug Dev. Ind. Pharm.*, 15, 2115, 1989.
58. Crawley, P., Drug-excipient interactions, *Pharm. Tech Eur.*, 13, 2001, 26.
59. Landin, M., Casalderrey, M., Martinez-Pacheco, R., Gómez-Amoza, J.L., Souto, C., Concheiro, A., and Rowe, R.C., Chemical stability of acetylsalicyclic acid in tablets prepared with different particle size fractions of a commercial brand of dicalcium phosphate dihydrate, *Int. J. Pharm.*, 123, 143, 1995.
60. Irwin, W.J. and Iqbal, M., Solid-state stability: the effect of grinding solvated excipients, *Int. J. Pharm.*, 75, 211, 1991.
61. Villalobos-Hernández, J.R. and Villafuerte-Robles, L., Effect of carrier excipient and processing on stability of indorenate hydrochloride/excipient mixtures, *Pharm. Dev. Tech.*, 6, 551, 2001.
62. Rowe, R.C., Sheskey, P.J., and Weller, P.J., *Handbook of Pharmaceutical Excipients*, 4th ed., Pharmaceutical Press, London, 2003.
63. Hancock, B.C., Colvin, J.T., Mullarney, M.P., and Zinchuk, A.V., The relative densities of pharmaceutical powders, blends dry granulation, and immediate release tablets, *Pharm. Technol.*, 4, 64, 2003.
64. Rowley, G., Quantifying electrostatic interactions in pharmaceutical powder systems, *Int. J. Pharm.*, 227, 47, 2001.
65. Hiestand, E.N. and Smith, D.P., Three indexes for characterizing the tableting performance of materials, *Adv. Ceramics*, 9, 47, 1984.
66. Roberts, R.J. and Rowe, R.C., The compaction of pharmaceutical and other model materials — a pragmatic approach, *Chem. Eng. Sci.*, 42, 903, 1987.

67. Amidon, G.L., Lennernas, H., Shah, V.P., and Crison, J.R., Theoretical basis for a biopharmaceutical drug classification: correlation of in vitro drug product dissolution and in vivo bioavailability, *Pharm. Res.*, 12, 413, 1995.
68. FDA, Guidance for industry: Dissolution testing of immediate release solid oral dosage forms, 1997.
69. Curatolo, W., Physical chemical properties of oral drug candidates in the discovery and exploratory setting, *Pharm. Sci. Tech. Today*, 1, 387, 1998.
70. Agoram, B., Woltosz, W.S., and Bolger, M.B., Predicting the impact of physiological and biochemical processes on oral drug bioavailability, *Adv. Drug Delivery Rev.*, 50, S41, 2001.
71. Kramer, S.F. and Flynn, G.L., Solubility of organic hydrochlorides, *J. Pharm. Sci.*, 61, 1896, 1972.
72. Serajuddin, A.T.M. and Pudipeddi, M., Salt selection strategies, in *Handbook of Pharmaceutical Salts: Properties, Selection, and Use*, Stahl, P.H. and Wermuth, C.G., Eds., Wiley-VCH, Weinheim, 2002, p. 135.
73. Johnson, K.C. and Swindell, A.C., Guidance in the setting of drug particle size specifications to minimize variability in absorption, *Pharm. Res.*, 13, 1795, 1996.
74. Pudipeddi, M., Serajuddin, A.T.M., Grant, D.J.W., and Stahl, P.H., Solubility and dissolution of weak acids, bases, and salts, in *Handbook of Pharmaceutical Salts: Properties, Selection, and Use*, Stahl, P.H. and Wermuth, C.G., Eds., Wiley-VCH, Weinheim, 2002, p. 19.
75. Dill, W.A., Kazenko, A., Wolf, L.M., and Glazko, A.J., Studies on 5,5′-diphenylhydantoin (Dilantin) in animals and man, *J. Pharmacol. Exp. Ther.*, 118, 270, 1956.
76. Serajuddin, A.T.M., Sheen, P.C., Mufson, D., Bernstein, D.F., and Augustine, M.A., Preformulation study of a poorly water soluble drug, alpha-pentyl-3-(2-quinolinylmethoxy) benzenemethanol: selection of the base for dosage form design, *J. Pharm. Sci.*, 75, 492, 1986.
77. Serajuddin, A.T.M., Sheen, P., Mufson, D., Bernstein, D.F., and Augustine, M.A., Physicochemical basis of increased bioavailability of a poorly water-soluble drug following oral administration as organic solutions, *J. Pharm. Sci.*, 77, 325, 1988.
78. Washington, N., Washington, C., and Wilson, C.G., *Physiological Pharmaceutics*, 2nd ed., Taylor and Francis, New York, 2001, chap. 5.
79. Davis, S.S., Hardy, J.G., and Fara, J.W., Transit of pharmaceutical dosage forms through the small intestine, *Gut*, 27, 886, 1986.
80. Smith, C.G., Poutisiaka, J.W., and Schreiber, E.C., Problems in predicting drug effects across species lines, *J. Int. Med. Res.*, 1, 489, 1973.
81. Chiou, W.L., Ma, C., Chung, S.M., Wu, T.C., and Jeong, H.Y., Similarity in the linear non-linear oral absorption of drugs between human and rat, *Int. J. Clin. Pharmacol. Ther.*, 38, 532, 2000.
82. Chiou, W.L., Jeong, H.Y., Chung, S.M., and Wu, T.C., Evaluation of using dog as an animal model to study the fraction of oral dose absorbed of 43 drugs in humans, *Pharm. Res.*, 17, 135, 2000.
83. Li, S., He, H., Parthiban, L., Yin, H., and Serajuddin, A.T.M., IV–IVC considerations in the development of immediate-release oral dosage form, *J. Pharm. Sci.*, 94, 1396, 2005.
84. FDA, Guidance for Industry: Extended Release Oral Dosage Forms: Development, Evaluation, and Application of *In Vitro/In Vivo* Correlation, 1997.
85. Dressman, J.B., Dissolution testing of immediate-release products and its application to forecasting in vivo performance, *Oral Drug Absorption: Prediction and Assessment*, Dressman, J.B. and Lennernas, H., Eds., Marcel Dekker, New York, 2000, p. 155.
86. Hariharan, M., Ganorkar, L.D., Amdion, G.E., Cavallo, A., Gatti, P., Hageman, M.J., Choo, I., Miller, J.L., and Shah, U.J., Reducing the time to develop and manufacture formulations for first oral dose in humans, *Pharm. Technol.*, October, 68, 2003.
87. Sistla, A., Sunga, A., Phung, K., Iparkar, A., and Shenoy, N., Powder-in-bottle formulation of SU011248. Enabling rapid progression into human clinical trials, *Drug Dev. Ind. Pharm.*, 30, 19, 2004.
88. Aubry, A.F., Sebastian, D., Hobson, T., Xu, J.Q., Rabel, S., Xie, M., and Gray, V., In-use testing of extemporaneously prepared suspension of second generation non-nucleoside reversed transcriptase inhibitors in support of Phase I clinical studies, *J. Pharm. Biomed. Anal.*, 23, 535, 2000.

89. Shah, V.P., Noory, A., Noory, C., McCullough, B., Clarke, S., Everett, R., Navkiasky, H., Srinivasan, B.N., Fortman, D., and Skelly, J.P., In vitro dissolution of sparingly water-soluble drug dosage forms, *Int. J. Pharm.*, 125, 99, 1995.
90. Abrahamsson, B. and Ungell, A.L., Biopharmaceutical support in formulation development, in *Pharmaceutical Preformulation and Formulation,* Gibson, M., Ed., CRC Press, Boca Raton, FL.
91. Waterman, K.C., Adami, R.C., Alsante, K.M., Hong, J., Landis, M.S., Lombardo, F., and Roberts, C.J., Stabilization of pharmaceuticals to oxidative degradation, *Pharm. Dev. Tech.*, 7, 1, 2002.
92. Al-Omari, M.M., Abelah, M.K., Badwan, A.A., and Aaber, A.M., Effect of drug matrix on the stability of enalapril maleate in tablet formulation, *J. Pharm. Biomed. Anal.*, 25, 893, 2001.
93. Shek, E., Ghani, M., and Jones, R.E., Simplex search in optimization of a capsule formulation, *J. Pharm. Sci.*, 69, 1135, 1980.
94. Bateman, S.D., Verlin, J., Russo, M., Guillot, M., and Laughlin, S.M., Development and validation of a capsule formulation knowledge-based system, *Pharm. Technol.*, 20, 174, 1996.
95. Shiromani, P.K. and Clair, J., Statistical comparison of high shear vs. low shear granulation using a common formulation, *Drug Dev. Ind. Pharm.*, 26, 357, 2000.
96. ICH Guidance for Industry Topic Q 1 A: Stability testing of New Drug Substances and Products, 2003.
97. Nash, R.A., Validation of pharmaceutical processes, in *Encyclopedia of Pharmaceutical Technology*, Swarbrick, J. and Boylan, J.C., Eds., Marcel Dekker, New York, 2002.
98. Pharmaceutical cGMPs for the 21st century: A risk based approach. US Food and Drug Administration, www.fda.gov/cder/gmp/index.htm
99. FDA Guidance to Industry: Bioavailability and bioequivalence studies for orally administered drug products — General considerations, 2002.
100. Cupissol, D., Bressolle, F., Adenis, L., Carmichael, J., and Romain, D., Evaluation of the bioequivalence of tablet and capsule formulations of granisetron in patients undergoing cytotoxic chemotherapy for malignant disease, *J. Pharm. Sci.*, 82, 1281, 1993.
101. Akers, M.J., Fites, A.L., and Robinson, R.L., Formulation design and development of parenteral suspensions, *Bull. Parent Drug Assoc.*, 41, 88, 1987.
102. Broadhead, J., Parenteral dosage forms, in *Pharmaceutical Preformulation and Formulation,* Gibson, M., Ed., Interpharm/CRC Press, New York, 2004.
103. Akers, M.J., *Parenteral Quality Control: Sterility, Pyrogen, Particulate, and Package Integrity Testing*, Marcel Dekker, New York, 1994.
104. Sweetana, S. and Akers, M.J., Solubility principles and practices for parenteral drug dosage form development, *J. Pharm. Sci. Technol.*, 50, 330, 1996.
105. Gupta, P.K., Patel, J.P., and Hahn, K.R., Evaluation of pain and irritation following local administration of parenteral formulations using the rat paw lick model, *J. Pharm. Sci. Technol.*, 48, 159, 1994.
106. Nema, S., Washkuhn, R.J., and Brendel, R.J., Excipients and their use in injectable products, *J. Pharm. Sci. Technol.*, 51, 166, 1997.
107. Strickley, R.J., Parenteral formulation of small molecules therapeutics marketed in the United States-Part III, *PDA J. Pharm. Sci. Tech.*, 54, 152, 2000.
108. Tang, X.L., Pikal, M.J., and Taylor, L.S., Design of freeze-drying processes for pharmaceuticals: practical advice, *Pharm. Res.*, 21, 191, 2004.
109. Puri, N., Pejaver, S.K., Parnajpe, P., and Sinko, P.J., Bioavailability of parenteral dosage forms, *US Pharmacist*, 26, HS41–HS53, HS55–HS56, 2001.
110. Suarez, S. and Hickey, A.J., Drug properties affecting aerosol behavior, *Respir. Care*, 45, 652, 2000.
111. Poochikian, G. and Bertha, C.M., Regulatory view on current issues pertaining to inhalation drug products, *Respiratory Drug Delivery VIII Conference*, Vol. 1, 2000, pp. 159–164 [www.rddonline.com].
112. Hickey, A.J. and Martonen, T.B., Behavior of hygroscopic pharmaceutical aerosols and the influence of hydrophobic additives, *Pharm. Res.*, 10, 1, 1993.
113. Hodges, C.C., Lloyd, P.M., Mufson, D., Rogers, D.D., and Wensley, M.J., Delivery of Aerosols Containing Small Particles through an Inhalation Route, US 6,682,716, January 27, 2004.

114. Rabinowitz, J.D., Wensley, M., Lloyd, P., Myers, D., Shen, W., Lu, A., Hodges, C., Hale, R., Mufson, D., and Zaffaroni, A., Fast onset medications through thermally generated aerosols, *J. Pharmacol. Exp. Ther.*, 309, 769, 2004.
115. Amiel, S.A. and Alberti, K.G., Inhaled insulin, *BMJ*, 328, 1215, 2004.
116. Zaffaroni, A., Therapeutic systems: the key to rational drug therapy, *Drug Metab. Rev.*, 8, 191, 1978.
117. Swintosky, J.V., Personal adventures in biopharmaceutical research during the 1953–1984 years, *Drug Intell. Clin. Pharm.*, 19, 265, 1985.
118. Allen, T.M. and Cullis, P.R., Drug delivery systems: entering the mainstream, *Science*, 303, 1818, 2004.
119. Novel Drug Delivery Systems (NDDS) Reports, Technology Catalysts International, Falls Church, VA, 2002.
120. Hastings, M.S., Li, S.K., Miller, D.J., Bernstein, P.S., and Mufson, D., Visulex™: advancing iontophoresis for effective noninvasive back-of-the-eye therapeutics, *Drug Delivery Technol.*, 4, 53, 2004.
121. Horspool, K.A. and Lipinski, C.A., Advancing new drug delivery concepts to gain the lead, *Drug Delivery Technol.*, 3, 34, 2003.
122. Meinzer, A., Mueller, E., and Vonderscher, J., Microemulsion — A suitable galenical approach for the absorption enhancement of low soluble compounds?, in *Absorption of Micellar Systems: Bulletin Technique Gattefossé,* Barthelemy, P., Ed., Gattefossé S.A., Saint-Priest, France, 1995, p. 21.
123. McClellan, M.B., Speech before PhRMA, March 28, 2003 (available at www.fda.gov).

Section II

Scientific Discoveries Application in New Drug Development

4

The Impact of Combinatorial Chemistry on Drug Discovery

Michael Rabinowitz and Nigel Shankley

CONTENTS

4.1 Introduction

4.1.1 Combinatorial Chemistry Will Transform the Pharmaceutical R&D Process

In the late 1980s major pharmaceutical companies, in the face of concern about dwindling new drug pipelines, initiated a major investment in a technology that promised to transform the efficiency of drug discovery. Combinatorial chemistry was to be employed to generate millions of compounds that could be screened against potential drug target proteins. The rapid deployment of the technology and re-tooling of major pharmaceutical research laboratories were fueled by the belief that the endless combination of molecules to form new compounds would significantly increase the probability of finding the next "blockbuster," the molecule with the best drug-like properties. The so-called high-throughput screening (HTS) (testing) of these molecules required equally revolutionary changes in biology laboratories. As the complexity of the human genome was unraveled, recombinant technologies came into their own, providing endless supplies of potential drug target protein in multiple engineered formats. Robots were introduced to process thousands of compounds through almost instantaneous readout cell- and protein-based assays. The investment, like the expectations generated, was enormous and significantly impacted the structure and composition of drug discovery scientific teams as they worked to harness the endless flow of new data. This chapter, prepared 15 years later, reviews the impact and development of combinatorial chemistry.

4.1.2 Industry Is Less Efficient than It Was

Between 1993 and 2003 the total U.S. pharmaceutical R&D spending was more than doubled with eight leading pharmaceutical companies investing approximately 20% of their sales revenue in R&D (Filmore, *Modern Drug Discovery*, 2004). However, during the same period the number of new medical entities (NMEs), defined in the analysis as drugs with a novel chemical structure, submitted for product approval to the Food and Drug Administration (FDA) fell from 44 in 1996 to 24 in 2003 (Figure 4.1). Similar trends were noted at regulatory agencies worldwide. The problem was highlighted in a white paper presented by the FDA in March 2004, which was bluntly entitled "Innovation or Stagnation" (http://www.fda.gov/oc/initiatives/criticalpath/whitepaper.html). In brief, the increased output of more effective, affordable, and safer therapies anticipated from the astonishing progress in basic biomedical science over the last two decades had not been realized.

The widely held view is that NMEs are invented in drug discovery programs, but that products are created from NMEs in the development process. Closer analysis of the data from the 1993 to 2003 period revealed that the number of NMEs entering early phase of clinical trials from drug discovery programs increased significantly (an 85% increase to Phase I and 90% increase to Phase II). In terms of these simple metrics, it could be argued that the productivity of research increased in line with investment and the basic science revolution. However, to suggest that the problem lies in the development process where the vast majority of NMEs failed to become products would be a gross oversimplification. NMEs primarily fail to become products because of shortcomings in safety or effectiveness, and both properties are co-owned by discovery and development. The high failure rate is proving costly to both the industry and patient. The disproportionate high cost of late-stage failures is being borne by fewer successful new products that drive higher pricing and predicate that companies restrict product development to large patient populations in

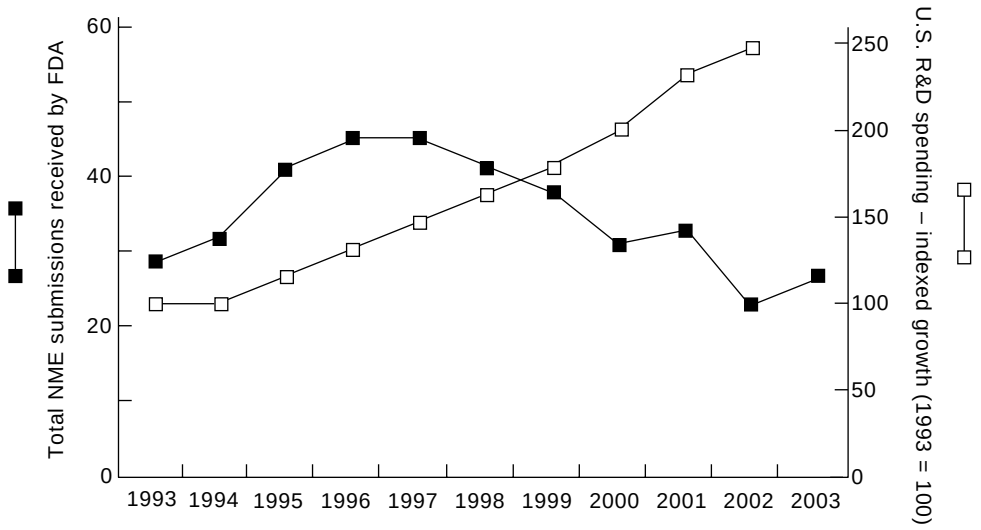


FIGURE 4.1
Ten-year trends in NME submissions to FDA and U.S. pharmaceutical research spending. The figure shows 10-year trends in U.S. pharmaceutical research and development (R&D) investment (*PAREXEL's Pharmaceutical R&D Statistical Sourcebook 2002/2003*) and the number of submissions to FDA of new molecular entities (NMEs), defined as drugs with novel chemical structure. (Redrawn from http://www.fda.gov/oc/initiatives/criticalpath/whitepaper.html.)

affluent societies. The prescription payers of these populations are in effect underwriting the decreased overall low productivity of the pharmaceutical industry.

The conclusion must be that the pharmaceutical companies are not selecting the right drug targets and patient populations or the NMEs are inadequate in terms of selectivity and biopharmaceutical properties. In this chapter, we review the impact of one of the technological breakthroughs — combinatorial chemistry — that governed significant changes over the last 15 to 20 years, not only in the initial discovery of NMEs, but also in the methods adopted for their all important characterization prior to entering the development process.

4.2 Background

4.2.1 The Traditional Interplay between Biology and Chemistry in the Drug Discovery Process

4.2.1.1 The Development of Medicinal Chemistry

Paracelsus (1493 to 1541) is credited with establishing the role of chemistry in medicine. It is evident in his oft-quoted, visionary view:

> "it is the task of chemistry to produce medicines for the treatment of disease since the vital functions of life are basically chemical in nature [.....] all things are poisons, for there is nothing without poisonous qualities. It is only the dose which makes a thing a poison."

However, until the development of analytical and synthetic chemistry in the 19th century, the discovery of new therapeutic substances remained focused on the biological investigation of natural substances usually secured from the apothecary garden. The new chemical techniques enabled the isolation and manufacture of pure substances. First, reproducible extraction procedures yielded individual active compounds from natural products and, subsequently, toward the end of the century, *de novo* synthesis was achieved of the simpler natural products such as aspirin, which was first marketed by Bayer in 1899. The next logical step was to systematically modify existing biologically active compounds, referred to as lead compounds, in an attempt to improve on their drug-like properties. By the 1930s the pharmaceutical industry had begun a sustained period of productivity, identifying lead compounds and transforming them into effective new therapeutic agents. These lead compounds came from a variety of sources including natural products, endogenous chemical messengers (hormones, neurotransmitters, enzyme substrates, etc.), random screening of the output of industrial chemistry processes (e.g., dyestuffs), and in the case of so-called me-too programs, simply from an existing drug. The latter programs were based on improving the biopharmaceutical properties or safety profile of existing drugs acting at a particular target or the optimization of a known side-effect that was perceived as potential therapeutic benefit in its own right. The success of the burgeoning pharmaceutical industry from 1930 through 1980 and the impact of the medicines produced on the quality and longevity of life are beyond question (Figure 4.2).

4.2.1.2 The Development of Bioassay

At the time of the first syntheses of the pharmaceutically active ingredients of natural products, the pioneers of pharmacology, Ehrlich, Langley, and Hill among others, were engaged in turning the study of the mode of action of drugs into a quantitative science. By

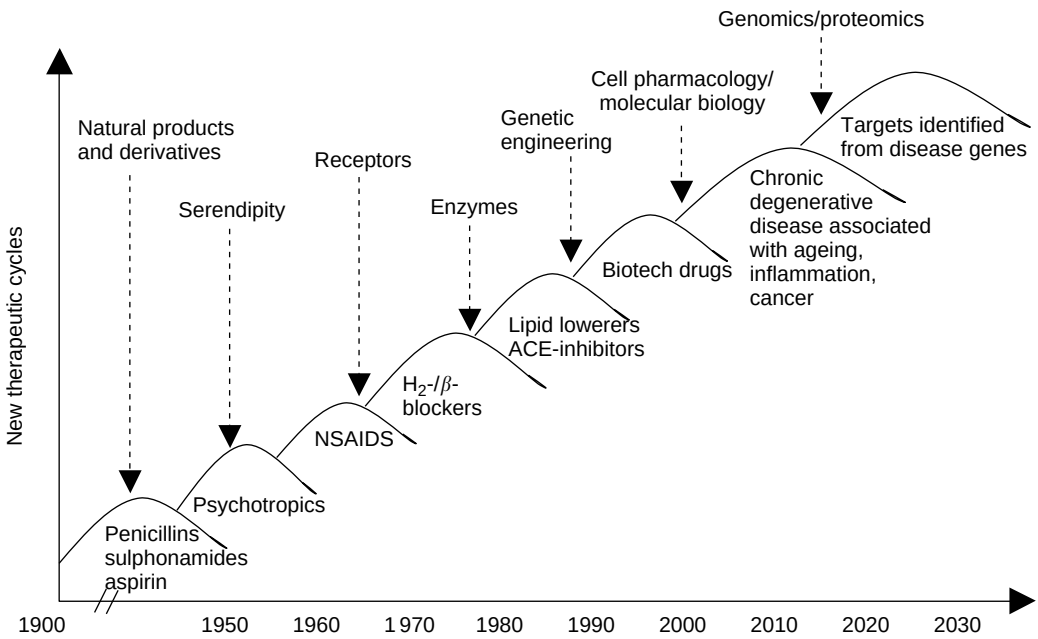


FIGURE 4.2
One hundred years of drug discovery technologies. (Redrawn from Adkins, S. et al., New Drug Discovery: Is Genomics Delivering? Lehman Brothers Research Reports, 1999.)

characterization of the relationships between concentration of active substance and the effects produced on biological systems and application of the newly established laws of thermodynamics governing the interaction between reacting chemical species, they inferred the existence of cell membrane receptors. These receptors were shown to be specific sites of action of the body's own chemical messenger systems, in addition to many of the natural substance-derived drugs being isolated and synthesized at the time. These early pharmacologists, and those like Clark, Gaddum, and Schild who subsequently built on their ideas, established the principles of bioassay that are still at the heart of the drug discovery process today.

Bioassays have been likened to analytical machines insofar as pharmacologists use them to assign biological properties to compounds in the same way a chemist measures the physical-chemical properties of molecules. If the fundamental role of the medicinal chemist is to optimize the pharmaceutical properties of so-called lead compounds by structural modification, then the role of the pharmacologist in the drug discovery process is to select, develop, and apply bioassays to provide relevant robust data that inform the medicinal chemist of the impact of the modifications he makes.

From the time of Ehrlich until the advent of readily available recombinant protein expression technologies in the 1980s, bioassay was predominantly based on the quantification of drug efficacy, safety, and distribution in living tissues or extracts thereof. Scientific reductionist drive provided increasingly more functionally simplified and homogenous assay systems. So, by the 1980s, techniques were available to study the action of new compounds on drug targets that, although originating from their natural expression environment, were present in experimentally amenable bioassays. Thus, the compounds could be studied in whole organisms, isolated intact organ or tissues, dispersed cell preparations, membrane homogenates, and ultimately, in the case of some targets, in enriched protein extracts. The pharmacologist now has at his disposal a full range of assays from the pure target through populations of patients to study the effects of compounds (Figure 4.3). Moving up the bioassay hierarchy from pure protein toward patient populations, it takes longer to obtain results and the comparisons of molecules that are increasingly complex as the compound activity is increasingly multifactorial. Only in the patient is the effect of the compound dependent on the expression of all the biopharmaceutical properties of the compound that the medicinal chemist seeks to optimize. Thus,

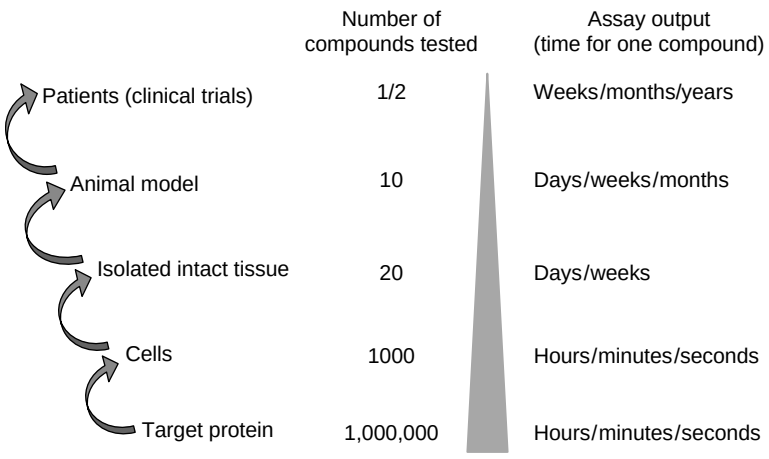


FIGURE 4.3
Bioassay compound throughput and assay timescales.

the assay readout is based not only on the primary activity at the target, but also on the adsorption from the site of absorption, distribution, metabolism, elimination, and any toxicity (ADMET), plus the natural variations in blood flow and chemistry and other homeostatic changes that occur in intact organisms.

4.2.1.3 The Choice of Bioassay: The Compound Progression Path

The challenge for drug discovery scientists is clear. Although the current medicinal chemistry-driven drug discovery process is often represented as a linear stepwise series of events (Figure 4.4), in practice, the process is holistic and iterative insofar as changes in structure impact all the properties of an NME (Figure 4.5). The simplest most functionally reduced assays usually involve the measurement of the primary interaction between the compound and its site of action. The assay readout from these basic or primary assays is usually a measure of the affinity of the compound for the drug target expressed in terms of the equilibrium dissociation constant, which is the concentration of the compound that occupies half the sites. For both the pharmacologist and medicinal chemist, these assays are perhaps the most satisfying. The assays are usually associated with low variance, and because the activity of the compounds can be expressed by the single chemical affinity parameter, the communication of progress of medicinal chemistry is straightforward. The functionally reduced assays are cheaper and faster, allowing the testing of many compounds in a time frame commensurate with the medicinal chemist's ability to conceive and synthesize new molecules based on the assay results. The problems begin to emerge when the chemist tries to address other properties of a potential NME–those governing the ADMET properties. These require the employment of the increasingly complex, expensive, and time-consuming intact physiological systems from

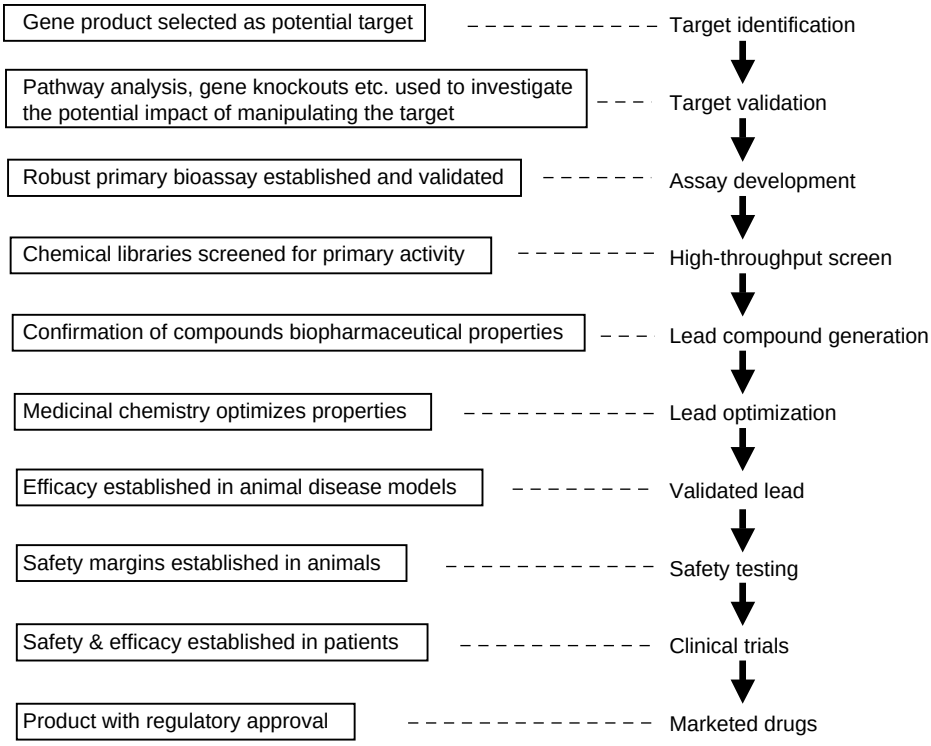


FIGURE 4.4
The drug discovery process — a stepwise view.

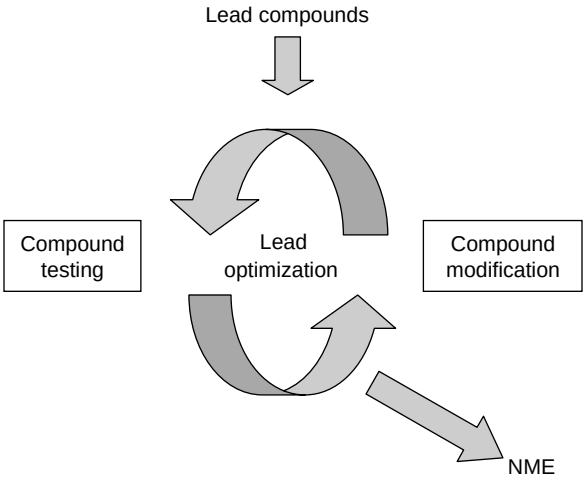


FIGURE 4.5
The drug discovery process — iteration between medicinal chemistry and bioassay.

which data cannot be obtained on each compound synthesized, and it certainly is never fast enough to feed the chemist's imagination and guide his synthetic choices. The situation was dramatically exacerbated by the implementation of combinatorial chemistry and massive increases in the number of new compounds that the chemist could produce for assay.

In an attempt to address this problem in the current drug discovery setting, a newly conceived program is usually reviewed against a set of project criteria with the aim of assessing the likelihood of success and defining the critical path a program should follow. Sir James Black has published his criteria based on 40 years' experience of working closely with medicinal chemists during which time he discovered both the first β-adrenoceptor and histamine H_2-receptor antagonists. These are reproduced in Table 4.1.

The two of Black's criteria that are most relevant to this chapter are: *Is a chemical starting point identified?* and *Are there relevant bioassays available to guide medicinal chemistry?*

4.2.1.4 Properties of Lead Compounds

The chemical starting point is the lead compound, which is usually defined along the lines of "a chemical entity that has known structure, high purity that reproducibly and predictably causes the desired biological effect; belongs to a chemical series that

TABLE 4.1
Criteria for Inception of a Medicinal Chemistry-Led Drug Discovery Project

1. Is there a basis for the desired selectivity of action — is the project purged of wishful thinking?
2. Is a chemical starting point identified?
3. Are relevant bioassays available to guide medicinal chemistry?
4. Will it be possible to confirm laboratory-defined specificity of action in humans?
5. Is there a clinical condition relevant to the specificity in point 4?
6. Is the project adequately resourced?
7. Does the project have a champion — someone with the necessary passion, conviction and energy?

Adapted from Black, J.W., in *Textbook of Receptor Pharmacology*, Foreman, J.C. and Johanson, T., Eds., CRC Press, Boca Raton, FL, 2003, pp. 271–279.

shows evidence of a structure–activity relationship (i.e., systematic structural changes lead to significant changes in activity) and has optimizable pharmacokinetics." In an attempt to rationalize the selection of a particular lead or lead series of compounds, attempts have been made to articulate these properties in terms of measurable parameters. For example, in 1997, Lipinski published an analysis of the properties of existing drugs to identify basic chemical characteristics that could be used as a filter to select "drugable" leads from screening hits. The properties considered usually include factors such as synthetic difficulty, analog potential, scale-up potential, purity and reproducibility, molecular weight, heteroatom/carbon ratio, lipophilicity, hydrogen bond donors and acceptors, number of rotatable bonds/rings, the absence of chemically reactive centers (i.e., no "toxic" features), and the ability to secure novel intellectual property.

4.2.1.5 The Compound Progression Path

With drug target set and lead compound(s) identified, the properties of an NME suitable for clinical evaluation can be specified more explicitly. These properties form the project goals and are written in terms of minimum acceptable values of the assay output parameters, as illustrated in Table 4.2. These criteria are then transformed into a compound progression path that provides the guidelines as to which compounds are progressed to more expensive, time-consuming assays (Figure 4.6). In practice, because of the iterative nature of the process, both the project goals and compound progression schemes are constantly updated as new information is generated both within the program and in the scientific community at large.

4.3 The Impact of Combinatorial Chemistry on Bioassay

Although the regulatory requirements for approval of a new drug have been subject to constant revision as tighter definitions of safety and efficacy are established, the fundamental properties required to be assigned to NMEs in the drug discovery process has not changed significantly over the last 30 years. Consequently, the overall approach to medicinal chemistry-driven programs outlined in the preceding sections has also remained largely the same. The impact of combinatorial chemistry and the simultaneous ability to produce effectively unlimited supplies of primary assay material for all the potential drug targets revealed by the human genome project can be viewed, in simplest terms, as a period of information overload. The introduction of robotic primary target activity assays allows the basic parameters to be obtained for all the compounds available, and without doubt increased the diversity of

TABLE 4.2

Specific Drug Discovery Project Goals for a Small Molecule NME

1. Specificity: defined mechanism of action, e.g., competitive receptor antagonist
2. Potent: high affinity for target, e.g., $pK_B < 8$
3. Selectivity: e.g., >100-fold selectivity over other pharmacological loci
4. Oral bioavailability: e.g., $\%F > 30\%$, $t_{1/2} > 4$ h
5. *In vivo* potency: e.g., <1 mg/kg for 50% maximum activity
6. Novel chemical structure with intellectual property rights protected
7. Scaleable synthetic route, reasonable cost of goods

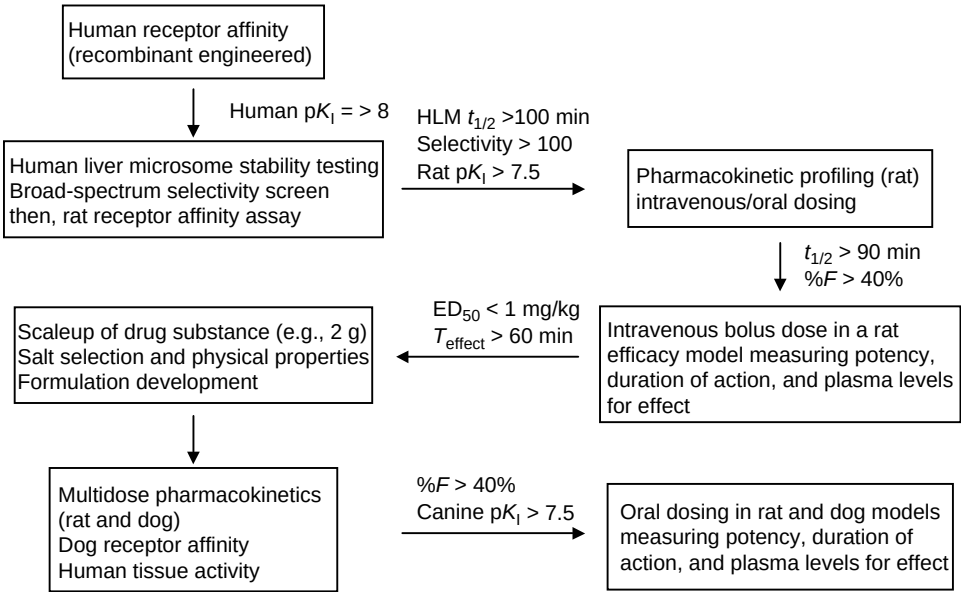


FIGURE 4.6
Receptor antagonist drug discovery program: compound progression scheme.

compounds that could be screened as potential leads in a program. The increased chemical diversity also allowed identification of chemical starting points for those targets where the traditional nonrandom screening sources of suitable leads were unavailable. Thus, the approach of adopting the natural ligand or substrate as the chemical starting point proved to be difficult in the case of many intracellular targets because the ligands tend to express relatively low affinity and selectivity. Similarly, screening of diverse libraries has provided the first small molecule ligands for many peptide hormone receptor targets for which working from the peptide hormone as a chemical template proved intractable. However, for all these successes in providing novel lead compounds, it soon became apparent that the lead optimization process would remain the rate-limiting step. A further worrying trend was that the new robotic laboratories for studying compounds at the protein or cellular level were often introduced at the expense of the systems assayist and his laboratory equipment. Indeed one unfortunate side effect of the molecular biological and combinatorial revolution has been the significant reduction in the advanced educational programs that provide the skilled scientists trained in performing the *in vivo* pharmacology experiments.

4.3.1 Combinatorial Chemistry

> Le mieux est l'ennemi du bien.
>
> —Voltaire, *Dictionnaire Philosophique*

With the accelerating advancement of reductionist biology in the 1980s and 1990s and with it, the advent of high-throughput target screening, came the need for an increase in the stores of compounds available for screening; not just in quantity of each but in sheer number — a concept closely linked with the notion of Chemical Diversity. To answer this call, chemists began to adapt methods long known in the synthesis of biopolymers for the

preparation of small molecules, methods that not only allowed for the rapid production of complex (and not-so-complex) structures for bioassay but also allowed the preparation of many diverse compounds in the process, often in a fraction of the time it would take to produce this diversity of structures by traditional solution synthesis methods. This process and set of techniques, as diverse and somewhat unrelated as they were, became to be known as combinatorial chemistry (Table 4.3). The following section briefly outlines the core concepts of combinatorial chemistry and attempts to place them into historical perspective now that the combinatorial chemistry "revolution" in industrial drug discovery is two decades old. This generational hindsight allows us the perspective to suggest what the real impact of these techniques has been in the pharmaceutical industry: The real revolution of combinatorial chemistry is the scientific codification and general acceptance of the fact that to make large numbers of compounds for bioanalysis, we have to loosen our requirements for overall analytical purity. And thus something was sacrificed in the process: we began thinking of compounds for screening as sets rather than individuals — allowing for the statistical treatment of purity of collections rather than of single compounds. The words of Voltaire are particularly relevant: "The enemy of the good is the

TABLE 4.3

Definitions of Combinatorial Chemistry Concepts

Combinatorial synthesis	The simultaneous preparation of all possible combinations of two or more mutually reactive sets of chemical monomers, either as individual compounds or mixtures of compounds
Library	A collection of compounds intended for bioassay that typically share a common pharmacophore or activity toward a given class of protein targets
Parallel synthesis	The simultaneous preparation of a subset of individual members of a combinatorial library
Solid-phase library	The use of polymeric, insoluble protecting groups in combinatorial or parallel synthesis that allow for the facile separation of target compounds from the rest of the reaction mixture under heterogeneous reaction conditions. Final target compounds may be assayed either while still attached to the resin beads or following cleavage
Solution-phase library	A library of compounds prepared under homogeneous reaction conditions without the aid of polymeric supports
Resin scavenging	The use of reactive polymer supports to selectively remove residual reagent or starting materials from a solution-phase mixture
Resin capture	The use of reactive polymer supports to selectively remove target products from a solution-phase mixture
Split and mix	The technique of preparing target compounds in a solid-phase library whereby resin-bound intermediates in the preparation of a solid-phase library are combined and then redivided for attachment of subsequent monomers. This technique produces mixtures of compounds but allows for the rapid production of very large compound libraries
Spatial addressability	The preparation of individual compounds in a parallel format such that the structural identity of each is known by its physical location in a matrix
Encoding	Defining the structural identity of a resin-bound compound in solid-phase library by marking the individual resin beads with a tag that is unique to the attached target compound
Deconvolution	The process of using iterative bioassay activity data to identify active components in a library mixture

perfect." The strictures of pristine compound purity adhered to fervently for decades in the medicinal chemistry laboratory were seen as roadblocks to the potential efficiency gains of combinatorial methods. By relaxing standards of chemical purity it was surmised that sheer numbers of compounds being created and screened would lead us down the path to quicker lead identification and optimization toward clinical candidacy. Now, 20 years on, it just may be that the enemy of the "good" has been found to be the "good enough."

Combinatorial chemistry is defined differently by many of its practitioners and reviewers. This ambiguity is probably to be expected for a field of endeavor that is still in its adolescence. For some, the term combinatorial chemistry refers to the technology involved in the creation of large diversity libraries. For these workers, this technology is often some version of resin-based "split-and-mix" diversity generation. For others, combinatorial chemistry is a broad term that refers to a collection of technologies and methods that allow the rapid production of all possible "combinations" of mutually reactive chemical fragments. In this latter definition, the exact method of library preparation (solid-phase, solution-phase, and solid-supported reagents, etc.) and the state of the target compounds upon completion of the effort (mixtures, discrete compounds, encoded resin-bound compounds, etc.) are incidental to the goal of the production of a true mathematical combinatorial assembly of similar chemical fragments. In practice, the term "combinatorial chemistry" refers to all technologies and applications used in the production of many compounds simultaneously, be they a true combinatorial collection, or a smaller subset of all possible combination of chemical fragments. Thus, in this definition, combinatorial chemistry also refers to multiple parallel syntheses (i.e., the synthesis of a row or column out of a full array) (Figure 4.7). In fact, what combinatorial chemistry has done for the bench chemist is to codify and thereby allow consensual adoption by those in the field of somewhat relaxed standards of analytical purity which we are willing to accept for biological assay data collection. This is an absolutely essential tradeoff to allow the production of new screening compounds to increase by orders of magnitude.

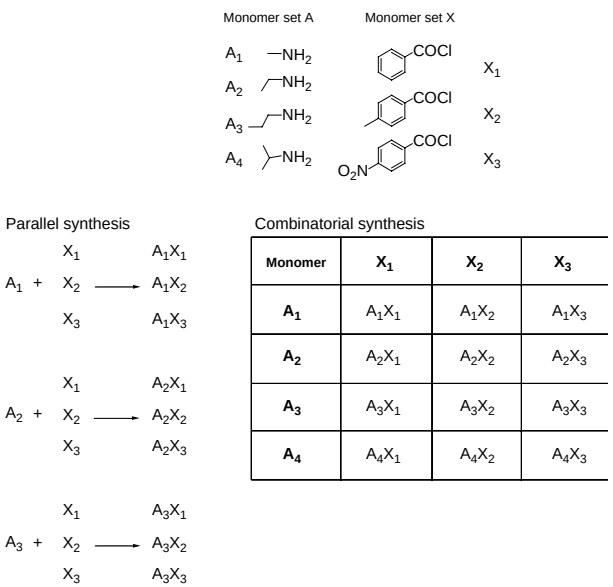


FIGURE 4.7
Traditional, parallel, and combinatorial syntheses.

4.3.1.1 High-Throughput Chemistry

Traditional corporate compound collections prior to the advent of HTS consisted of thousands or tens of thousands of discrete entities, usually in the form of analytically pure powders. Additionally, it was not unusual that these samples were available in relatively large quantities of 0.1 g to >1 g. It became clear in the late 1980s that the capacity to screen compounds for biological activity would greatly outstrip the compound collections of most pharma companies. Industry leaders thus started on the inevitable track of building their compound collection through purchasing of compound sets and by investing heavily in the technology of combinatorial chemistry.

At that time, the basic technologies of combinatorial chemistry had been around for about a decade but had found application mainly in protein and oligonucleotide synthesis with the preparation of "libraries" of compounds, i.e., collections of analogs prepared using common chemistry performed simultaneously. These researchers built upon the Nobel Prize winning work of Merrifield in solid-phase peptide synthesis by showing that not just one but many molecular entities could be prepared through a logical sequence of chemical reactions involving a controlled and divergent polymerization. The scientific underpinnings of combinatorial chemistry, as practiced today, have changed little from 1985. What has changed is the technology available to enable the science. What follows is a brief accounting of the most important concepts of combinatorial chemistry that have had the most profound impact on the way that drug discovery research is done in the pharmaceutical setting.

4.3.2 Mixtures, Libraries, and Compound Arrays

4.3.2.1 Discrete Compound Synthesis

Traditional medicinal chemistry was built upon the solid foundation of traditional organic synthesis, whereby single compounds were produced in relatively large (ca. 1 g) batches that were easily crystallized to high purity. The techniques of combinatorial chemistry shifted this tradition toward the more modern idea of increasing compound "throughput" through the preparation of many more compounds in smaller amounts (nanograms to milligrams) in purities that, as mentioned above, were measured as statistical collections rather than individuals (i.e., average purities and the number that pass minimal purity cut-offs).

As first practiced by Geysen and Houghton, the preparation of combinatorial libraries produced discrete compounds of known identity through a technique known as "spatial separation," which simply means that individual compounds in the library are produced discretely and are not mixtures. Such spatially addressable compound sets are produced in such a way as to keep separate the reaction flasks or resin beads containing the individual components of the library and perform bioassays on the discrete compounds, one at a time. Thus, if the "history" of the reaction conditions performed in each flask or on each solid support, the identity of the compounds produced is known, without resort to structure elucidation techniques. Initially, this technique, after typically an extensive reaction development stage, allowed the preparation of between 10 and 1000 discrete combinatorial products.

4.3.2.2 Mixtures: Split-and-Mix Synthesis, Deconvolution, and Encoding

Classical Drug Discovery focused on the biological evaluation of individual purified compounds. But in light of HTS and combinatorial chemistry, it became widely accepted that the biological screening of mixtures of compounds should, in theory, provide much more biological data in the same amount of time used for bioassay of individual compounds,

thus increasing efficiency and reducing costs. It was reasoned, after all, that the screening of mixtures had precedent in the biological evaluation of natural products (leading to the discovery of penicillin, amongst other drugs), and that in most cases, biological activity expressed in a mixture of compounds would be either orthogonal or simply additive thus facilitating the identification of active components.

While this may in fact be the case for natural product mixtures, it is rarely the case when dealing with synthesized mixtures. Despite our attempts to create real molecular diversity in the test tube, our efforts have not even begun to anticipate the true diversity of atomic connectivity within "drug space" (estimated to be of the order of 10^{63} unique compounds, theory, famously in this case, greatly outpacing the amount of matter in the universe). Thus, combinatorial chemistry was never practically able to produce true chemical diversity and compounds produced in such library format ended up looking very much like one another, with the attendant similarities in biological activity profiles.

It is very straightforward to produce mixtures of compounds and much of the history of Organic Chemistry has involved the avoidance of such endeavors. In theory, one simply must mix all of the monomeric parts of a final compound together, either simultaneously or consecutively, and mixture will be obtained. But performed in this manner, the mixture would most likely not be assayable (i.e., it would result in meaningless data that cannot be interpreted in terms of compound structure). However, one of the triumphs of combinatorial chemistry is that it has freed the Drug Discovery scientist from the strictures of "one compound, one assay" by showing that to some degree that the preparation of compound mixtures can result in a more rapid collection of structure activity relationship (SAR) data if performed in a controlled and logical manner. The most important technique for doing this is so-called "split-and-mix." Indeed, this is the only practical way to produce and handle very large libraries of compounds. Houghten has reported the preparation of 10^{12} decapeptides using split-mix techniques.

In this solid-phase technique, resin beads are split into equal portions and each portion is allowed to react with one each of the first set of monomers. The resin is then recombined, mixed thoroughly, prepared for reaction with the second set of monomer (typically involving a de-protection step), and again split into a number of portions equal to the number of reactants present in the second group of monomers. This process is repeated until all of the monomer sets have been ligated to the growing resin-bound analytes, and, rather than mixing the final resin portions together again, they are kept separate so that knowledge of the last chemical operation each batch of resin had seen was retained. Thus, for each of the cleaved mixtures, the pools have in common the identity of the final monomers, and it is therefore the identities of the monomers making up the remaining positions in the final active compounds that remain unknown.

An additional complication of the bioassay of mixtures, be they natural extracts or combinatorial in nature, is that any activity present in a pool of compounds must still be assigned to the individual components to render the screening exercise useful for Drug Discovery. This is more straightforward with respect to combinatorial mixtures due to the fact that, by definition, the structural type is known, and what is not known is just the specifics of the attachment of monomer. Nonetheless, structure elucidation of active components must be made and this has traditionally been done through what is referred to as "deconvolution."

Deconvolution may be either iterative in nature or positional. Iterative deconvolution involves the systematic resynthesis of subsets of active mixtures, gradually narrowing down the activity in an iterative resynthesis or screening cycle until a single active component has been identified. There are a number of iterative deconvolution strategies that will often, if performed carefully and on ideally behaved mixtures, yield the same results. It often begins with identifying an active pool of compounds through bioassay, and if this pool was prepared using ordinary split-and-mix techniques, then the identity of the final

monomer in the reaction sequence is known. Then, smaller mixtures of compounds, each bearing that final "active" monomer, can be prepared and assayed for activity. Once again, the most active pool, having been prepared by the split-and-mix strategy, now reveals, in addition to the identity of the final monomer, the identity of a second monomer of the active final compound. This process is shown schematically in Figure 4.8. This process of split-and-mix followed by iterative deconvolution is perhaps what most people refer to as "combinatorial chemistry" itself, and dominated the literature and the venture capital of the field a decade and more ago and promised to provide us with a faster route to identifying the drugs that were already out there just waiting to be discovered! The reality of the technique was somewhat more sobering. Impurities present in these poorly characterized mixtures sometimes were discovered to contain all of the observed biological activity (the same may also be said about the poorly characterized "pure" compound). It was not uncommon for companies to retain entire groups of "resynthesis" chemists to help track down the activity of a given sample, occasionally having to prepare and purify divined impurities after resynthesis of the desired combinatorial compound revealed no biological activity.

A second popular deconvolution strategy is known as positional scanning. In this technique, multiple libraries containing the same compounds, but in different mixtures, are produced and assayed in a single round of synthesis and screening. In practice, by synthesizing one full library for each monomeric position, with each individual library pool defining, via spatial separation, the identity of a single monomer, the "code" of the most active compound can be read from the identity of active pools in each full library. The obvious drawback for positional scanning is the necessity to prepare "*n*" libraries containing the same compounds, where *n* is the number of degrees of freedom or points of diversity. Thus, in practice, the combination of the uncertainties of deconvolution (you can never truly be

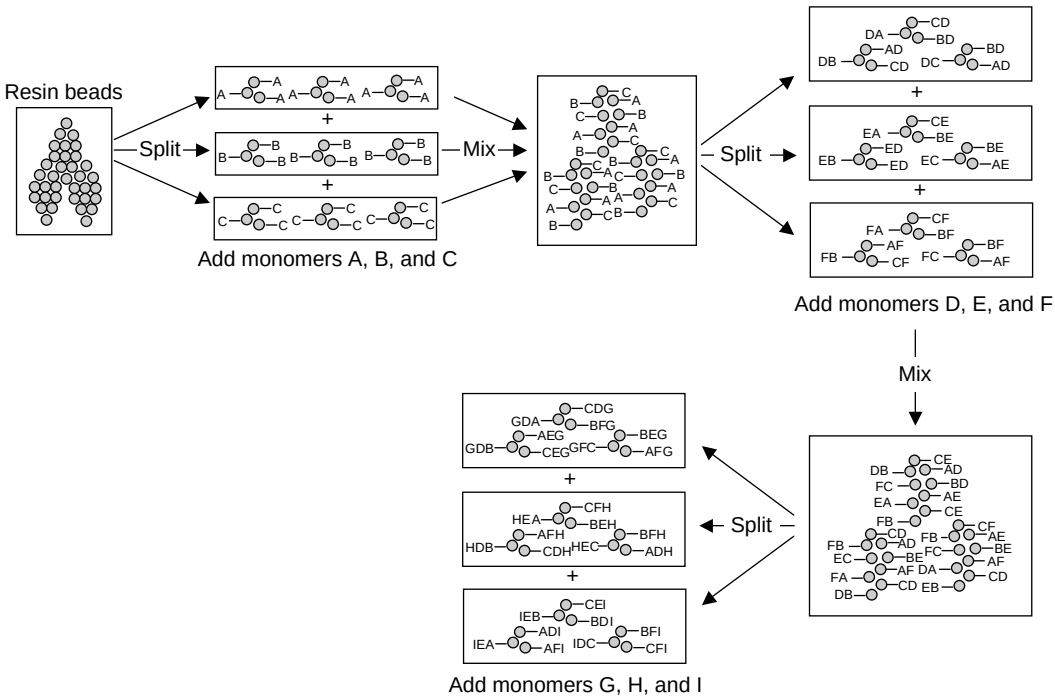


FIGURE 4.8
The split-and-mix technique for the production of combinatorial mixtures.

sure that the most active compound in a library gets identified — so-called hard pooling) coupled with the fact that nondiverse libraries will have many compounds of varying activities present in mixtures leading to departure from ideal behavior.

4.3.2.3 Encoded Combinatorial Synthesis

A technique that was very popular but had diminished in significance over the past 5 years is the technique of encoding. In theory, this should be a very useful and concise method for the structure elucidation of active compounds within a resin-bound mixture. In its simplest form, encoding involves the orthogonal placement of chemical identification tags on resin beads containing the target compound, with the identity of the individual monomers that make up the final product being "encoded" by the very constitution of the tags themselves. This technique has proven fairly tricky because it requires that discrete compounds prepared on individual beads in mixtures be assayed either while still attached to the bead of origin or else cleaved from that single isolated bead and assayed while retaining the bead itself for analysis. Why is this so? Because the orthogonal synthesis strategies for analyte and tag leave the tag still attached to the bead, ready for analysis should the analyte prove to have biological activity. Another technological hurdle for the encoding technique has been the obvious requirement that the tag be much easier to analyze and identify than the test compound itself (otherwise one could simply use some analytical technique to identify the structure of the product rather than an inactive chemical surrogate). This has been done most successfully by a number of techniques including secondary amide strategy, mass encoding, polyhalogenated aromatics, and radio-frequency tagging. The fact that encoding techniques for compound identification in the split-and-mix protocol is so specialized and often requiring extraordinary hardware is punctuated by the fact that whole companies or divisions have been set up or repositioned to produce, promote, and capitalize the individual encoding strategies.

It may be conjectured at this point that the vast majority of expenditure by the pharmaceutical and biopharmaceutical industries on the "combinatorial chemistry revolution" has been in the areas of split-and-mix synthesis, deconvolution, and encoding or decoding and much novel and innovative equipment designed to facilitate these techniques. Ultimately, while these techniques produced a great number of compound libraries for evaluation in the drug discovery setting, often only nanograms of final compound — quickly depleted after only a few rounds of screening — were procured. Furthermore, the production of minute quantities of compounds outstripped our understanding of the optimal methods of storing them. The "good enough" solution was to store these compounds in solubilized form. The ubiquitous solvent DMSO seemed to provide the perfect answer; miscible with water in all proportions and typically innocuous to biological systems in low concentrations. But the repeated freeze or thaw cycles of this kind of storage wreaked havoc on solubility and chemical stability causing the loss of countless compounds throughout industry.

Combinatorial chemistry as described above, fraught with problems of expense, purity, compound identification, and storage have largely been supplanted in the modern medicinal chemistry laboratory by the more measured and reliable techniques broadly termed "parallel synthesis."

4.3.2.4 Parallel Synthesis

Parallel synthesis is the term loosely applied to the preparation of discrete compounds in the format of a spatially addressable array. It is often used in contrast to "mixture-based" combinatorial synthesis. While it is true that the term "parallel synthesis" may encompass the preparation of a fully combinatorial multidimensional array of compounds, in practice, it most

typically refers to the preparation of discrete compounds in a one- or two-dimensional array with the aim of elucidating the SAR around a specific chemical scaffold. In this respect, the definition of Baldwin is useful: "In brief, a 'parallel synthesis' strategy essentially requires the synthesis of a highly reactive intermediate via a series of simple steps, then its subsequent reactions with a number of different reagents…" Parallel analog syntheses are typically used to prepare 10^1 to 10^3 compounds at a time. Much automation has become available to facilitate parallel library preparation, including microwave heating, robotic reagent handling, 96-well filter plate reaction formats, automated parallel purification, fluorous methods of purification, resin-bound reagents, resin scavengers, and others. In essence, though, parallel synthesis has at its core more traditional organic chemistry techniques of reactivity, handling, and purification. More so than all other techniques of the "toolbox" of combinatorial chemistry, parallel synthesis has come closest to reaching the ideal imagined decades ago. There is barely a medicinal chemist working today who does not use parallel techniques in one form or another for hit-to-lead and lead optimization work. It truly can be said that of all the advancements that the field of combinatorial chemistry has brought to the world of Drug Discovery, parallel synthesis techniques have been the most pervasive and the most widely accepted by the pharmaceutical industry. Not yet realized is the dream that every working medicinal chemist would "…not be surprised when asked to prepare a couple hundred thousand compounds for the screening project next week." Combinatorial chemistry has led to the development and acceptance of some very useful tools to assist in the more efficient discovery of drugs, but it has replaced none of the traditional disciplines. The process of Drug Discovery was for a while slower and costlier than it was before the advent of combinatorial chemistry and HTS. These newer techniques have proven useful and are starting to show some promise in terms of the efficient discovery of new medicines for unmet medical needs. However, they have fallen short of being the magic bullets that were supposed to identify unique drug molecules for each assayable drug target.

4.3.3 Types of Library Design Strategies and Their Uses

4.3.3.1 Diversity Library

One of the early ideals of combinatorial chemistry was the belief that scientists could use these technologies to enable the unprecedented preparation of millions of compounds (both peptides and small molecules) simultaneously. The resulting very large libraries would be called "diversity" or "random" libraries and were (and still are) designed to fill as much chemical space as is allowable by the combinatorial connectivities of hydrogen, carbon, nitrogen, oxygen, sulfur, phosphorous, and halogen atoms. It became quickly apparent that true diversity covering all of chemical space, even within the confines of "drug-likeness" is an unattainable goal. It has been estimated that within the confines of protein space alone, the number of possible unique 245 amino-acid peptides derived from the 20 most common amino acids at just one molecule for each library member, the resulting mass would far exceed that of the universe. So with the acceptance that absolute chemical diversity is unattainable, workers set about simply trying to fill specific chemical space with as diverse sets of compound scaffolds and monomers as practicable (usually meaning commercially available). As practiced then, filling diverse chemical space through the production of nominal diversity libraries for the pharmaceutical industry ended up taking the form of producing large collections of compounds that ultimately were either poorly represented in a given corporate compound collection, or considered to represent a pharmacophore that traditionally showed great promise in lead identification through screening (so-called privileged structures). But due to the number of compounds produced in such a library, and to the effort involved in developing, optimizing,

and producing the library itself, and analyzing the average purities (quality control), these libraries are rarely produced for a single screening target. They are meant to add to the quality and diversity of screening sets in order to increase the chances of finding drug-like hits when little structural information about the protein target or ligand pharmacophores is known.

In the initial enthusiasm to create such random libraries, it was not unusual in earlier days to worry less about the properties of the final compounds synthesized in these libraries, and more about ease of chemical synthesis. It can be conjectured that at that time, reactive orthogonality of chemical building blocks was more of a design concern than was drug-likeness of the target molecule. Thus, many large libraries (and small) were produced from aryl chlorides, arylisocyanates, simple amines, alcohols, and carboxylic acid derivatives. Not only did they end up being fairly undrug-like, but also repetitive and pointlessly nondiverse. After all, it was not a very straightforward thing to have hydroxyl, mercapto, and amino groups in final compounds that were prepared using acid chloride and isocyanates. Consequently, for a time, combinatorial synthesis of polyaromatic amides free from hydrophilic functionality became very much in vogue. The place of the derived target molecules in the world of known drug molecules with respect to size and physicochemical properties was tenuous.

In recent years, the maturing of the field of diversity library production has led to the concept of "diversity-oriented synthesis." Researchers working in this field have accepted the early shortcomings of diversity-based library approaches (*vide supra*) and adapted the preparation of fairly complex molecules, containing multiple nonaromatic ring systems, hydrophilic functionality, and multiple controlled stereo centers, to solid-phase techniques, thus giving new life to what many had come to view as a tarnished field. While typically no more "drug-like" than earlier diversity library output, these new diversity-oriented libraries have been able to produce products with utility in disrupting complex protein–protein interactions and helping to elucidate cell-signaling pathways.

4.3.3.2 Focused Libraries

The second important kind of chemical library that has been produced frequently in the drug discovery setting is the so-called focused library, a name chosen to distinguish it from the random library mentioned above. Typically such a library is prepared when there exists some degree of structural or pharmacophoric information about a desired target–ligand interaction. The distinction here is that some small molecule or peptide lead molecule has already been identified and validated (e.g., from the literature, HTS of more diverse compound collections, or analogy to known ligands of similar targets). The goal of the focused library is to further refine this initial information to determine if there exists a true SAR in the lead series, if there is evidence of optimizable pharmacokinetics in the series, if structural variation can lead to novel intellectual property and so forth. These libraries are commonly designed around a single protein target and are typically of the order of 10 to 1000 compounds in size. They may involve mixtures but are far more commonly prepared using spatially addressable parallel array techniques. These libraries are by far the most common types produced today in the drug discovery setting, and they often play vital roles in hit-to-lead and lead optimization programs. They have taken a leading role at the forefront of combinatorial chemistry as it is currently practiced in industry, because:

1. They produce a manageable number of compounds that can all be analyzed for identity and purity.

2. They rely on proven, robust and commonly available techniques, and instrumentation that are now quite prevalent in the medicinal chemistry laboratory and in the repertoire of the medicinal chemist.
3. Their components are produced typically in a spatially addressable format requiring no deconvolution or encoding strategies for identification of active components making for very quick analysis of SAR.
4. They are, as currently practiced, nearly always directed by chemoinformatics, i.e., the application of drug-like product information in library design and selection of monomers.

For these reasons, the production of small focused libraries has largely surpassed larger "diversity" libraries in prominence in both academia and industry (*vide infra*). Most large pharmaceutical companies today have refocused their large, technology-driven combinatorial chemistry laboratories of yesteryears into smaller groups aimed at the production of these focused library sets for specific biological targets. These new "hit-to-lead" or "technology-enabled synthesis" groups, while founded in the split-and-mix, encoded, and diversity library trend of early combinatorial chemistry, are now typically pared-down chemistry units who employ the simplest, most efficient HTS (solution or solid phase) and purification technology with the goal of rapidly producing tens to hundreds of milligrams each of individually purified compounds for full spectroscopic characterization and bioassay. The intended missions of such groups are to help validate a novel lead series or elucidate preliminary SAR around a screening hit.

It is clear then by this measure that the technological advances of combinatorial chemistry play a role as facilitators of more traditional medicinal chemistry in the pharmaceutical setting, advancing the efficiency of traditional organic synthesis such that chemical series destined for failure as drugs can be identified and eliminated more rapidly than before. Thus, we now find ourselves at a time beyond the adolescence of combinatorial chemistry, well into its adult years where, in the words of Lester Mitscher, "After an initial euphoric period when many investigators thought that any novel compound had a realistic chance of becoming a drug, realism has now returned and libraries are being constructed that reflect the accumulated wisdom of the field of medicinal chemistry."

4.4 Combinatorial Chemistry in the Modern Drug Discovery Setting — Commentary

"Boy, the food at this place is really terrible." "Yeah, I know … and such small portions."

—Old Vaudeville joke

This joke, used to great effect in the opening lines of the film "Annie Hall," is emblematic of the pharmaceutical industry in the early and mid-1990s. The answer, it seemed a decade ago, to weak pipelines and waning efficiency (R&D investment vs. sales) was to make more, because new medicines were just waiting to be "discovered." Thus, the faster a company could synthesize and screen new candidate molecules, the faster that company would find them (and keep competitors from finding them first).

4.4.1 The Promise

"The overwhelming benefits of combinatorial chemistry are that it is faster, more efficient, and cheaper than traditional chemistry and can result in the synthesis of literally hundreds or thousands or even millions of compounds in a very short time." Yes it can, but one may ask, "To what end?" It is estimated that the drug industry in the United States has invested over the past decade $35 billion on research and development, a sum that is over twice that of the previous decade. And in that same period, the number of new molecular entities (NMEs; distinguished from line extensions, combination therapies, and reformulations) has halved to about 40. This is despite the record increases in the number of compounds currently in corporate R&D pipelines (early drug discovery to Phase II clinical trials). Thus, as an industry over the past decade, we have invested heavily in equipment, facilities, and staff, and have participated in a revamping of the overall drug discovery process from a modular system of basic research followed by preclinical development, followed by full development. We have gathered, using the tools of HTS and combinatorial chemistry, a vast collection of information about new targets and exciting new ligands for those targets. Yet we have not, to this point, been successful in turning all of this information into safe and effective medications for novel treatments for unmet medical needs. It is sometimes said that the research arm of the R&D process does not make drugs — we collect data. Ideally, drug researchers collect enough data to support the hypothesis that a given pure substance, if prepared on a scale and formulated and dosed correctly, will have a beneficial effect for patients suffering from a particular disease. It is the development arm of R&D that actually prepares the drug and is charged with demonstrating its safety and efficacy. What is clear about the current state of drug research is that through the aegis of modern genomics, DNA microarrays, HTS, and combinatorial chemistry, we have collected unprecedented amounts of information about new targets and novel chemical structures that modulate them. We have invented numerous important chemical tools for the elucidation of biological pathways, and have seen an exponential increase in the number of chemical entities available for bioassay. But we have failed to date in converting this vast wealth of information into new drugs.

It has been estimated from a 1996 survey using traditional medicinal chemistry techniques, a chemist could provide on average one new compound per week at a cost of $7500 per compound. Furthermore, it was also estimated that by employing the technologies of combinatorial chemistry, that same medicinal chemist could produce over 800 compounds per week at a cost of $12 per compound, realizing a cost efficiency of over 600-fold, and the production of nearly three orders of magnitude more compounds in the same period of time. It is clear then why the promise that this new technology held for drug discovery was so alluring and spawned so many new companies a decade or more ago. Of course, very few of those companies exist in their original forms today, having disappeared altogether, or having switched their focuses from those of technology based (so-called broadly enabling platform technologies) to product based, reflecting a complete change in the marketplace and in the flow of venture capital. Allowing that these 1996 estimates of cost and time savings are even remotely accurate, it cannot be said in 2005 that the majority of chemists in the pharmaceutical industry produce 800 compounds per week on annual average. Thus, what becomes immediately obvious is that an increase in total number of compounds prepared and bioassayed over the last 10 to 15 years has not yet resulted in increases in the *real* measure of efficiency: Novel and improved *drugs* reaching the patient. The early period of the combinatorial chemistry revolution has taught us one thing above all else: The absence of high-quality compounds in lead discovery and optimization cannot be offset by an increase in number. More compounds were prepared and screened in the same period of time for far less money per compound compared with prior

periods, but our industry has not yet seen any clear return on the initial investment and increased "efficiencies." Still today there is advocacy of the tenuous argument that an increase in compounds synthesized and screened is an end unto itself (and by implication, directly proportional to efficiency in the discovery of new drugs).

While the idea of students routinely producing hundreds of thousands of compounds per week has not come to pass, another prediction that combinatorial chemistry will become absolutely routine technology in the workplace and no longer an arcane technology relegated to specialists, is certainly correct, but in a way probably not envisioned earlier: Through the use of automated techniques to speed the interplay of traditional synthesis/pharmacological evaluation, not replace it. We have come to the realization that indeed, "combinatorial chemistry excels at speed, but machines lack insight."

4.4.2 The Reality

The current view of combinatorial chemistry is that it has resulted in a series of technologies that have greatly altered the way in which pharmaceutical drug discovery is undertaken compared to the situation prevalent just a decade ago. By creating new technologies for the rapid assembly of collections of new molecules for biological evaluation, and by codifying new standards for the acceptance of purity—the purity of collections rather than the absolute purity of individual compounds—the technologies inspired by combinatorial chemistry have irreversibly positioned themselves as indispensable in the end-to-end process of drug discovery. Yet it has not lived up to the initial promise of replacing the science of drug discovery by fully automating the complex and iterative interplay of chemical synthesis and biological evaluation. In retrospect, this should come as a surprise to no one. At its core, combinatorial chemistry is a set of technologies, not unlike recombinant protein assays, high-field NMR, molecular modeling, robotics, etc., that greatly support the central science of drug discovery, but by no means has replaced it in the pharmaceutical industry. It certainly is not to be considered a fad, except possibly by those ultimately disappointed investors a decade ago who backed technology over product. This premise is getting greater acceptance as the initial hype of the science of combinatorial chemistry and the notion that true "chemical diversity" would be at the fingertips of all in the drug discovery setting (viewed now as "unrealistic") replaced by the reality of the utility of the technologies that it embraces. Even the notion of academia of the future routinely training chemists in the preparation of hundreds of thousands of compounds in a week has been tempered by the realization that chemical careers in industry are made by the ability to solve problems rather than by simply training for the skills of a given technology. This view is strongly supported by industrial chemists and academicians alike (see Further Reading). Knowledge of combinatorial chemistry emerges as a useful skill-set for a prospective employee to have, but that alone is not sufficient to secure a position in the pharmaceutical industry.

To understand better how the role or vision of combinatorial chemistry in drug discovery has changed over time, it is instructive to understand how the nature of combinatorial libraries has changed. Roland Dolle, formerly of Pharmacopeia, Inc., a company involved in the commercialization of combinatorial technologies and combinatorial libraries, has published annually since 1998, a compendium of chemical libraries that have been published in the chemical and patent literature from both industrial and academic sources. In the initial survey of publications from the years 1992 to 1997, nearly 100 combinatorial libraries were published and reviewed with over 10% of them consisting of at least 1,000,000 members, and nearly 30% of them containing over 10,000 members (Figure 4.9). Libraries consisting of fewer than 100 members made up less than 20% of the total. In 2003, nearly 400 combinatorial libraries were reported, a greater than four-fold increase over the number published in

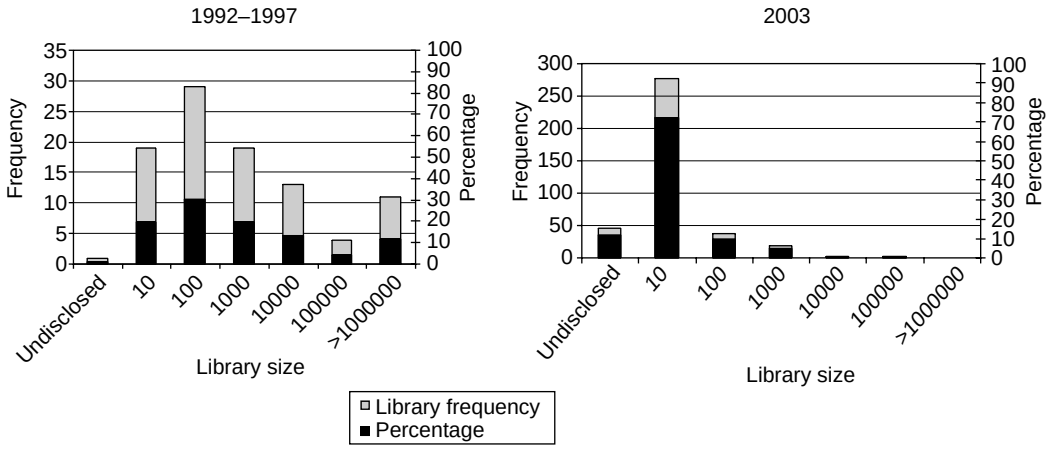


FIGURE 4.9
Trends in compound library size.

the 6-year period 1992 to 1997, but the intended use of the libraries had clearly changed: Less than 1% of these libraries contained over 1,000,000 members while more than 70% contained under 100 examples; particularly note the number of synthesized libraries without disclosure of the number of compounds produced (1% or 1/96 for 1992 to 1997, and 12% or 45/385 for 2003). The trend then is clear: Researchers working in the field of combinatorial chemistry emphasize less on total numbers of compounds produced and more on novelty of chemotype (fewer peptides; more small molecules) and advances in technology (i.e., solid-phase synthesis, resin scavengers, solution phase techniques, etc.).

This trend has been described by Hermkens and Müller in an independent analysis who observed that over the last 10 years, library size has dramatically decreased with an increased emphasis on purity of individual compounds rather than statistical purities of compound sets. Importantly, split-and-mix approaches to the preparation of large mixtures have given way to parallel techniques. These authors also point out what has become obvious to all those working in the field: That synthetic expediency should never dictate the structure of targets to be prepared in library format. Simply making many compounds is not a viable counterargument to this precept. The halcyon call of "chemical diversity" has now been replaced by the concept of "chemical similarity" to drug-like molecules. This has been a direct result of the realization that hit rates for combinatorial diversity libraries have been comparatively very low (vs. more traditional screening of the amassed corporate compound collections) as well as the observations that successful drug molecules occupy a very small sector of chemical space. Furthermore, it is now appreciated that successful drugs also appear, more often than not in a very narrowly defined window of physicochemical parameters. This recognition, most famously codified by Lipinski and coworkers, has prompted the refocusing of combinatorial drug discovery efforts within industry to take into consideration issues of hydrophobicity, molecular weight, hydrogen bonding, and number of aromatic rings.

The realization that not all atomic connectivities of C, O, N, S, P, and halogens that lead to interesting protein ligands will lead to viable drug candidates is perhaps the single most important correction in the steady advancement of combinatorial chemistry in the drug discovery setting. This topic has been amply reviewed for those interested in the details. This wholesale change in the way that industry and many in academia now view the significance of combinatorial chemistry in the drug discovery setting is a direct consequence of the realization that the global success rate for the discovery of new medicines has fallen

in the past two decades rather than increased. An anticipated increase in efficiency (as measured by NMEs generated per dollar spent) was generally accepted to be a direct consequence of our investment in automation in chemical synthesis and in bioassaying. The actual reduction in efficiency has been directly discussed in a number of articles and most frequently takes the form of investor concerns about pharmaceutical pipelines. The backlash has been severe by investors and scientists alike.

Venture capital at the moment largely avoids finding its way toward start-ups who bill themselves as "technology-based." There is an increased emphasis on products rather than "platform technologies." The basis for this is clear. Platform technologies in the field of combinatorial chemistry have either been incorporated as tools in the workplace or have disappeared from the scene entirely. Many companies that started off in this field have either gone out of business, been purchased (for their technology and libraries) by larger firms, or have transformed themselves into "biopharmaceutical" companies with active collaborations in NME discovery.

Scientists have also adopted a less ebullient and more measured view of the real efficiencies that combinatorial chemistry has to offer. It is not uncommon to hear the terms "unrealistic," "not particularly useful," and "flavor of the month" when medicinal chemists and drug discovery managers talk about the initial promise that an increase in speed and throughput were sufficient *per se* to bolster the collective drug pipelines of the pharmaceutical industry. Particularly compelling has been the understanding that advancements in automation are advancements in technology, and that drug discovery has always been and remains a science. As a science, drug discovery at its core involves the very intimate dialog between chemistry and pharmacology, and medicine with multiple nuanced iterations between these disciplines. There is no one formula that will be applicable to all drug discovery programs. So, as asked in a recent article, "Has combinatorial chemistry's fifteen minutes of fame come to an end?" The answer clearly is "No" — it simply has matured beyond the initial hype into a generally useful set of techniques that help speed us toward lead discovery and allow us to get more quickly to the still-problematic steps in the drug discovery process, as clearly described by Ed Scolnick, former head of R&D at Merck: "But the rate-limiting step of drug discovery is not finding a lead, it's turning a lead into a drug … Solving the problems of absorption, specificity, metabolism, and safety is only slightly easier today — if any easier — than it was 10 years ago."

Further Reading

Adang, A.E.P. and Hermkens, P.H.H., The contribution of combinatorial chemistry to lead generation: an interim analysis, *Curr. Med. Chem.*, 2001, 8, 985–998.

Adkins, S., Hetzel, C., Smith, I., Walton, J., Butler, C.A., Hazlett, R., and Hayashi, I., New Drug Discovery: Is Genomics Delivering? Lehman Brothers Research Reports, 1999.

Derra, S., Combichem Excels at Speed, but Machines Lack Insight, in http:// www.dddmag.com/ShowPR.aspx?PUBCODE=016&ACCT=1600000100&ISSUE=0409&RELTYPE=PR&ProdCode=00000000&PRODLETT=C 2004.

Dolle, R., Comprehensive survey of chemical libraries yielding enzyme inhibitors, receptor agonists and antagonists, and other biologically active agents: 1992 through 1997, *Mol. Diversity*, 1998, 3, 199–233.

Dolle, R.E., Comprehensive survey of combinatorial library synthesis: 2003, *J. Comb. Chem.*, 2004, 6, 624–679.

Dolle, R.E. and Nelson, K.H., Jr., Comprehensive survey of combinatorial library synthesis: 1998, *J. Comb. Chem.*, 1999, 1, 235–282.

Editor Scolnick quote: *Financial Times*, March 21, 2000.
Filmore, D., Thayer, A.M., and Willis, R.C., Pipeline challenges, *Mod. Drug Disc.*, 2004, 7, 28–34.
Frost and Sullivan Product News. Analysis of the Combinatorial Chemistry Markets, 2004 (available at: http://www.laboratorytalk.com/news/fro/fro176.html)
Furka, A., Combinatorial chemistry: 20 years on, *Drug Disc. Today*, 2002, 7, 1–4.
Geysen, H.M., Meloen, R.H., and Barteling , S.J., Use of peptide synthesis to probe viral antigens for epitopes to a resolution of a single amino acid, *Proc. Natl. Acad. Sci. U.S.A.*, 1984, 81, 3998–4002.
Gilman, V., The right combination, *Chem. Eng. News*, 2004, May 3, 45–49.
Hermkens, P.H.H. and Müller, G., The impact of combinatorial chemistry on drug discovery, in *Ernst Schering Research Foundation Workshop*; *Small Molecule-Protein Interactions*, Vol. 42, 2003, pp. 201–220.
Houghten, R.A., Parallel array and mixture-based synthetic combinatorial chemistry: tools for the next millennium, *Annu. Rev. Pharmacol. Toxicol.*, 2000, 40, 273–282.
Konings, D.A.M., Wyatt, J.R., Ecker, D.J., and Freier, S.M., Deconvolution of combinatorial libraries for drug discovery: theoretical comparison of pooling strategies, *J. Med. Chem.*, 1996, 39, 2710–2719.
Lebl, M., Parallel personal comments on "Classical" papers in combinatorial chemistry, *J. Comb. Chem.*, 1999, 1, 3–24.
Lee, A. and Breitenbucher, J.G., The impact of combinatorial chemistry on drug discovery, *Curr. Opin. Drug Disc. Dev.*, 2003, 6, 494–508.
Lipinski, C.A., Lombardo, F., Dominy, B.W., and Feency, P.J., Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, *Adv. Drug Del. Rev.*, 2001, 46, 3–26.
Mitscher, L.A. and Dutta, A., Combinatorial chemistry and multiple parallel synthesis, in *Burger's Medicinal Chemistry and Drug Discovery*, Abraham, D.J., Ed., 6th ed., Vol. 2, *Drug Development*, Wiley, New York, 2003.
Pesidis, A., Combinatorial chemistry, *Nat. Biotech.*, 1998, 16, 691–693.
Pirrung, M.C., Spatially addressable combinatorial libraries, *Chem. Rev.*, 1997, 97, 473–488.
Pirrung, M.C., Chau, J.H.-L., and Chen, J., Indexed combinatorial libraries: nonoligomeric chemical diversity for the discovery of novel enzyme inhibitors, in *Combinatorial Chemistry*, Wilson, S.R. and Czarnik, A.W., Eds., Wiley, New York, 1997, pp. 191–206.
Professor J. Baldwin laboratory website: http://www.chem.ox.ac.uk/jebaldwin/research/parallel_synthesis.htm
Reymond, J.-L., Outrunning the bear, *Angew. Chem. Int. Ed.*, 2004, 43, 5577–5579.
Sanchez-Martin, R.M., Mitoo, S., and Bradley, M., The impact of combinatorial methodologies on medicinal chemistry, *Curr. Topics Med. Chem.*, 2004, 4, 653–699.
Wilson-Lingardso, L., Davis, P.W., Ecker, D.J., Hébert, N., Acevedo, O., Sprankle, K., Brennan, T., Schwarcz, L., Freier, S.M., and Wyatt, J.R., Deconvolution of combinatorial libraries for drug discovery: experimental comparison of pooling strategies, *J. Med. Chem.*, 1996, 39, 2720–2726.
Young, S.S. and Ge, N., Design of diversity and focused combinatorial libraries in drug discovery, *Curr. Opin. Drug Disc. Dev.*, 2004, 7, 318–324.

5

High-Throughput Screening: Enabling and Influencing the Process of Drug Discovery

Carol Ann Homon and Richard M. Nelson

CONTENTS

5.1 Introduction

Drug discovery must be a dynamic process so that it can incorporate the latest scientific discoveries and state-of-the-art technologies. Inherent in this process is the screening of samples for their potential to be drugs. Early drug discovery used animals to test these samples, usually natural products (NPs) or chemically modified compounds using the natural product structure as a scaffold for their effects. These early NPs and their analogs formed the basis for most of the modern pharmaceutical companies. As the pharmaceutical industry progressed into the early 1950s, many companies, as well as the National Cancer Institute, were using cellular assays to screen NP extracts from microbes and plants to look for new anticancer and antibacterial drugs. This NP extract screening continued to grow until by the mid-1980s; there was a need to screen thousands of extracts a week. This level of screening required a faster rate of testing because the stability of these test samples was unknown. As the pharmaceutical companies grew, they were establishing compound libraries from all of their drug-discovery programs. These compound libraries were literally the company's gold

reserves because many of the compounds were proprietary and often patented. As patents on chemical matter expire after 20 years, these archived compounds would be starting points for the synthesis of new compounds that could be protected by patent. Compounds from drug-discovery programs had usually been tested in animals; therefore, much was known about their pharmacological activity and any potential safety issues. By the late 1960s and early 1970s, it was a common practice to screen a sample bank of both NP extracts and selected compounds from the compound library against isolated proteins being explored as drug targets.[1] The limited amounts of these proteins that could be isolated by conventional methods often limited the number of screens that could be run. However, major advances in biotechnology, such as cloning and polymerase chain reaction (PCR), helped to overcome this limitation. Cell lines can now be used as factories in which recombinant proteins are overexpressed to produce as much as is needed by any screen. Thus, the components, samples, and targets for the modern age of screening came together. In the late 1980s, the term high-throughput screening (HTS) came into use to describe this rapid turnaround testing of samples against targeted proteins. The term itself originated within Pfizer (personal communication). The HTS process uses automation to produce the faster testing rate and to reduce the amount of repetitive work performed by humans. As the term HTS describes a process change associated with the maturation of several scientific discoveries and advances in technology, it is not defined by any particular level of throughput, but rather by the concept that more can be tested by applying suitable advanced technology to the process. The process continues to evolve, producing higher throughputs and applying improved scientific thinking to the process. The compound libraries have been dramatically enhanced by the addition of millions of compounds produced by combinatorial chemistry (CC) to add diversity to the collection as well as to increase coverage of chemical space. The number of targets for drug screening is still being defined, but has grown considerably with the sequencing of the human genome. HTS is practiced by most, if not all, pharmaceutical companies as well as by many of the larger biotechnology companies.[2] It is clear that random screening of compound libraries is really a rational approach to discover what may be the starting point for a new drug program. A successful screening program for drug discovery requires the screening of several targeted proteins to determine which target is the best for a full lead optimization program in a therapeutic area. In the last 10 to 15 years, HTS has changed in many ways but has stayed with its initial objective to produce more results faster. The process, as practiced at the major pharmaceutical companies, has required companies to build a complex infrastructure over the years to support it. HTS uses the very latest technologies as well as sophisticated, production-level automation for assay performance and protein production. There is a careful balance between selecting new technologies that add value to the process while not falling trap to ones that simply consume resources (cutting edge vs. the bleeding edge). The process is carefully scheduled and orchestrated to meet timelines and coordinate the upstream supply chain and output. The level of output is determined by the need for a predefined number of leads per year from the HTS process according to calculated attrition rates. Many organizations resemble a research factory because of the expected-output scheduling process that is needed to sustain itself. Today, the major driving force of HTS is to generate high-quality data that can be used as the basis of the drug development phases that follow HTS, starting with hit-to-lead (HTL), through lead optimization (LO), to the development candidate. New assay technologies combined with new automation equipment and quality control procedures ensure this higher-quality data. HTS produces data sets of tens of millions of data points each year that have required new data-analysis tools for their proper evaluation. All of these improvements and tools have enabled the HTS process to produce a better product. The demonstration that high-quality data could be generated by automated assays led to a growth in the use of automated assays for HTL support, for LO support, and for adsorption, distribution, metabolism, excretion, and

toxicity (ADMET) assays. This chapter will discuss the infrastructure, key components, and common practices of a typical HTS operation in a major pharmaceutical company. Operations such as this can support a successful drug-discovery program and put drugs in the clinic and on the market.[3–5]

5.2 High-Throughput Screening Today: The High-Quality, Automated Assay

High-throughput screening, in the simplest of terms, is the assaying of samples for activity on biological targets with the help of automation; but the level by which HTS does this requires special data-handling tools to process the raw data, to assess the quality of the data, and to archive the data into databases where it can be easily searched. Every HTS screening assay is carefully built from a large toolbox of components that have been established within the screening group. The selection of the biological targets is usually done by the therapeutic areas, so that is not normally within the realm of the HTS operation. However, it is critical that the HTS group be involved with target selection as consultants to ensure that selected targets can be screened and to ensure that the best assay for the screen will be developed. Not every assay, even an excellent bench assay, will be able to stand up to the rigors of today's HTS screen, so many screens use assay technology that has been developed intentionally for HTS (Sections 5.2.3 and 5.2.4). The best HTS screen is an assay that has been built intentionally to be a screen, starting with bench performance and moving to the automated performance in stages. This chapter will describe the components of the HTS function including automation, the sample bank, assay technology, and data-analysis tools.

5.2.1 Automation

A typical HTS laboratory today may produce more than 10^5 data points per day, each and every day of the week. At this level of throughput, it is essential that the HTS function uses dependable, robust automation. This automation must also be highly versatile to perform the different assay procedures optimally, which often requires the use of different devices in different assays. At the simplest level, automated devices and instruments are available for most steps of the assay. These devices include automated pipettors, reagent addition (broadcast) devices, plate washers, and plate readers. The selection of a particular device for use in HTS applications should be based on its performance under HTS-like testing conditions. Most devices are brought into the laboratory for extensive testing under a broad range of test conditions that mimic the types of procedures that would be used in an HTS assay. This testing looks at robustness under the throughput requirements of HTS. The demands of HTS may require minor modifications of the device such as the upgrading of the pump units supplied with the plate washers to handle the higher volumes and extended periods of operation.

One very important device is the plate reader, which can be rate limiting in HTS. Most laboratories use multimodal readers that can detect various forms of fluorescence as well as luminescence and absorbance. The traditional readers are photomultiplier-based devices that usually read from one well to the next. This process can take considerable time for 384-well and higher-density plates. A more desirable HTS reader type images the entire plate with a charge-coupled device (CCD) camera. The latter device is usually a faster reader for 384-well and higher-density plates. Imagers can capture significant cross talk from one well to another, but with proper set up, they can produce data of equal quality.

Two principal imagers used in HTS laboratories are the ViewLux (PerkinElmer Life Sciences, Wellesley, MA, U.S.A.) and the Leadseeker (GE Healthcare/Amersham Biosciences, Chalfont St. Giles, Bucks, U.K.). The performance in any particular mode of these multimodal readers can vary from one instrument to another, so most HTS laboratories have at least two different readers available for assay development and HTS.

As the HTS process was maturing, the microplate evolved from the 96-well standard of the late 1980s and early 1990s to the 384-well standard by the late 1990s to the incorporation of low-volume 384-well or 1536-well plates that are being used by most groups today. The automated devices must be able to handle all of these formats and allow for easy crossover from one format to another. The device's automation can be the simple operation of the device by pushing the go button or by programmed processes, or it can be the inclusion of the device into an automated workstation where plates will be handled in a batch mode. These methods require either an external computer or built-in processors. Examples of workstations include the addition of automated stackers to a device or the addition of a small, robotic arm that loads and unloads plates to and from the device. There is no universal stacker that allows the movement of stacked plates from one device to another. Companies are trying to come up with a stacker that will work with all of their own instruments, but none has yet been released. For workstations, an operator is responsible for the on-time movement of stacks of plates from one device to another. All these automated processes require the proper identification of the plate that is being processed. Today, most plates are identified by means of a barcode affixed to a side of the plate. In the future, two-dimensional (2-D) barcodes, which contain quite a bit more information as well as redundancy, may be used on microplates. The 2-D codes are presently used mainly on many different types of microtubes. A small laboratory can process stacks of plates through several assay steps via workstations with the help of people power quite well. However, people can be variable when handling many steps in an assay simultaneously. When the processing throughput exceeds 100 plates in a day, a higher level of automation is usually installed to perform the tasks.

Several companies build full, robotic systems that process at least 100 plates in a 20 to 24 h period. Only three companies specifically built systems to meet the challenge of ultra-high-throughput screening (ultraHTS/uHTS/UHTS). In 1994, this was defined as the testing of 100,000 unique sample wells per day,[6] which translated at the time to the screening of 1200 microplates in a 96-well format in 24 h. The first systems that achieved this rate were delivered by Zymark (now Caliper, Hopkinton, MA, U.S.A.) in 1998 and by Zeiss (Jenna, Germany) in 1999. Both systems were built as a series of integrated workstations, called *modules*, that use robots or conveyor belts to move plates from one workstation or device to another.[7–9] These systems are built on assembly line concepts and use different configurations to handle the various steps of the different assay procedures. Based on this assembly line idea, the fastest throughput for any system will be achieved if it is used in a unidirectional manner because this means that no device will be used more than once in an assay procedure. Thus, the transit time of the first plate is a sum of the time per workstation plus the sum of the incubation times. However, the next plate typically follows a minute or two later, and so forth, for the entire run. The typical processing time of a workstation is about 1 min for the 96-well plate and can be less than 2 min for a 384-well plate. Therefore, if a homogeneous assay (no separation steps) is performed with five modules and a 30-min incubation time, more than 200,000 samples can be tested in 20 h. Although the systems were originally built to handle the 96-well plates, both were easily upgraded to handle both 384-well and 1536-well formats by changing the devices within certain modules. The third system was built by Aurora Biosciences, who coined the term ultraHTS system (UHTSS).[6] This system, planned in 1994 around proprietary technology, includes an extremely high-density plate with 3456 wells — the nanoplate. The volume of

these wells required assays with very small volumes of 1 to 2 μL. The major driver in the early 1990s to go to ultraHTS was the influx of compounds expected from CC efforts that would expand compound libraries to well over a million compounds. These systems are presently used primarily for assays in which the extreme miniaturization is indicated because of high expense, such as the β-lactamase technology.[10,11] This technology was originally proprietary to Aurora as well and is now available from Invitrogen (Carlsbad, CA, U.S.A.). The cost of using the latest technology can be high, so there has been a constant effort to reduce assay volume as a way to save on reagent costs. The miniaturization of screening assays will be discussed further in Section 5.2.5.

Early laboratory robots were unreliable, but today, these systems perform quite well. Today's robots simply move plates from one robot-friendly position to another, such as the entrance pad of a plate reader. These simplified movements combined with the low weight of a plate allow engineering to simplify the robot designs. As seen in industrial application of robots, robots that are defined and used for a specific application will work day in and day out quite well. It is always best to keep the automation as simple as possible to get the highest level of performance. This is usually accomplished by minimizing the number of moveable parts associated with the automation. Stackers have also become more reliable. This was due, in part, to the standardization of the microplate by an effort of the Society for Biomolecular Screening (Danbury, CT, U.S.A.) in association with the American National Standards Institute (ANSI, Washington, DC, U.S.A.), but also due to the use of simpler stacker mechanisms. Today, there are many choices for devices, workstations, and fully automated systems. The selection as to which automated devices to purchase for HTS should be driven by a clear set of specifications that define the use of the automation. The choices can be expensive, and therefore, replacement may not be possible, so it is important to choose well.

A critical need to perform any automated assay is an automated liquid handler that can add a reagent of the assay or the test sample or both, and mix the contents well after addition. When choosing this device, the specification needs to include all of the potential uses: the volume range of the pipetting steps, all the plate formats to be used, the requirement for fixed or disposable tips or both, and the requirement for the change-out of disposable tips or the washing of fixed tips. The volume range requirement should allow for calibration over a defined range or calibration at a specific volume. The calibration will be different for different liquid classes. The volume delivered must be both accurate and precise, within specifications (usually a CV of less than 5%). The best devices allow for handling different plate formats either by changing the pipettor head or the array. The importance of calibration using specific assay reagents cannot be overstressed. To meet all of these requirements, a laboratory may be required to have several different devices for different types of pipetting needs. Indeed, in our HTS laboratory, we have several different pipettors as well as different devices for broadcast reagent delivery. Broadcast reagent delivery devices are used to add the same reagent to all wells of the microplate. Protein and other reagents can stick to different surfaces, so the HTS laboratory must have several pipettor options — such as plastic tips, glass syringes, or silicone tubing — to add these reagents accurately. The assay must dictate the automation that will be used, and not the reverse, if the assay is to produce high-quality data. An HTS group must have all the required tools for all different types of assays, but the rewards are in the data.

HTS requires that tens of thousands of compound solutions be ready for use on the day of screening. These HTS-ready plates are made from stock compound-solution plates (called *mother plates*) by pipetting aliquots from the mother into many daughter plates. Automated 96- or 384-tip pipettors are used to stamp out these replicate plates rapidly because the solvent is dimethyl sulfoxide (DMSO), which is hydroscopic (Figure 5.1). Mother plates are aliquotted with disposable, polypropylene tips to prevent cross

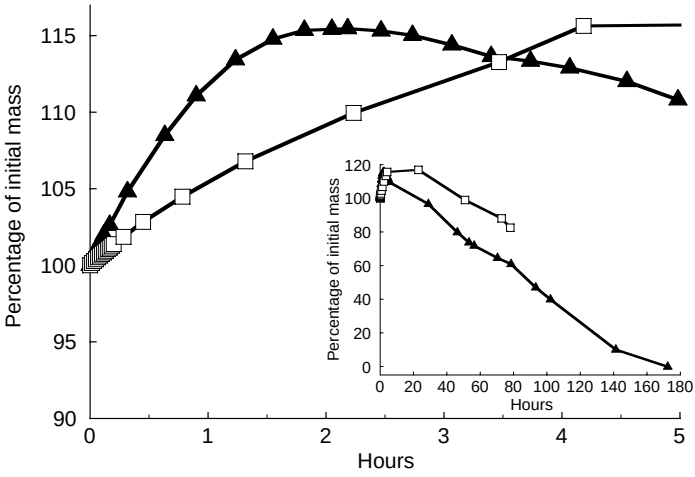


FIGURE 5.1
Time-dependent mass change of low volumes of DMSO under ambient laboratory conditions. Dry DMSO was pipetted (2 or 5 μL) into the wells in alternating rows of a low-volume 384-well microplate, and the plate was incubated on the laboratory bench under ambient conditions (approximately 21°C and 35% relative humidity) and weighed periodically using an analytical balance. Filled triangles are data from wells initiated with 2-μL DMSO; open squares are data from wells initiated with 5-μL DMSO. Inset shows data from the extended time course.

contamination from one mother plate to another. Ideally, the same pipettor device can perform both 96 and 384 operations, but some pipettors require a head change to handle this. A large laboratory will usually not change heads but will have two instruments with assigned heads to increase efficiency of the processing from mother to daughters. There are several major manufacturers of instruments that offer 96- and 384-disposable-tip options. The instruments can pipette as low as 2 μL into a dry plate with both accuracy and precision. These same instruments are used to perform HTS assays as well. For HTS performance, the pipettor may add a small robotic arm to move plates around the deck positions. This converts the pipettor device into a workstation. Some use stackers to reach workstation status. Stackers work well for dispensary operations in which small batches are usually processed but are not efficient for assay use if larger numbers of plates are to be processed because of the back-and-forth motion of the plates.

5.2.2 The Sample Library

At the start of the modern era of screening, the compound library at the larger pharmaceutical companies consisted of their particular group of compounds, which had been made over the past 50 years.[1] Some companies had a NP extract library as well to add to their sample bank. Companies wanted to quickly add diversity to their compound library so they worked with various external companies and academic laboratories to acquire novel structures. The acquired compounds were either new chemical scaffolds or they were scaffolds that were not well represented in the library. The external acquisition programs continued for many years. When the use of CC became widespread, the need for external acquisitions decreased. CC was seen as the means to add both depth and diversity to the collection (see Chapter 4 for further discussion of combinatorial chemistry). A consequence of the addition of hundreds of thousands of CC compounds was a decreased interest in the NP extracts.[12] Although a hit in an extract indicates a very potent compound structure, it can be quite difficult to find the specific structure within the complex mixture that is demonstrating activity. NP purification and isolation require a special skill set and can be a very slow process. However, NP hits may actually be the drug candidate and, thus, the whole LO process is avoided. As NP and the CC programs are both quite expensive, many companies had to choose one or the other, and most decided that CC would fulfill their needs. Many of the NP extract programs have been terminated. CC promised to provide a rapid path forward from the hit because each hit would have several analogs within the library to provide instant structure–activity relationships (SAR) and established synthetic routes. The first CC libraries

were produced with many compounds within one well. The compounds could be at different concentrations because of variable yields, which were often much lower than planned. At Boehringer Ingelheim Pharmaceuticals, Inc. (BIPI), we found that known hits were often missed in these complex mixtures, and the deconvolution of these mixtures to find the hit compound was time-consuming. The next generation of CC compounds produced a single compound per well, but the compounds were neither purified nor quantitated. Both of these early experiments with the use of CC cost many hours of effort to resolve the true hits from artifacts such as unreacted reagents. Most companies today produce much smaller sets of compounds with what is called automated or high-throughput chemistry. Each of the compounds is analyzed for purity, and the companies are slowly bringing the purity criteria up to the same level as single compounds produced by traditional medicinal chemistry, usually 90% or greater. The compounds are delivered to a company's dispensary unit in 96 formatted microtubes with a known quantity of compound. Wells with low yields or poor purity are culled. To expedite these final steps, microtubes with 2-D barcodes can be handled by automated workstations that remove the tubes with unsuitable samples and then reconfigure the remaining tubes. The first CC efforts also produced some very large compounds with molecular weights over 500 and sometimes as high as 1500. It is now recognized that these molecules should be more lead-like, which means lower molecular weights and less-complicated structures that are more soluble.[13] Automated chemistry continues to evolve, and the final chapter in this story has not been written.

Combichem libraries and NP extracts can be viewed as extreme examples of mixture samples. In the early days of HTS, companies often created defined compound mixtures from their compound libraries to reduce the cost of screening and the workload. These mixtures usually contained from five to ten known compounds. If the mixture hit in an assay, the sample was deconvoluted back to the individual compounds to find the single active component. However, an active mixture did not always lead to a specific active compound. Companies tried to avoid the deconvolution step by including each compound in two unique mixtures. This matrix method requires both mixtures to be active to define the active compound.[14] A problem arose with this approach when the HTS laboratories did not throw out mixture hits even when the confirming mixture did not hit because of their concern that the active compound was masked by the components of the second mixture. The laboratories then tested the contents of both mixtures as individual compounds. Thus, the costs and the workload were not dramatically reduced compared with screening single compounds. Over time, it became evident that although mixtures could be used to find active compounds, there was not much in the way of savings, and there were always concerns that active compounds were being missed. In addition, the data generated by testing mixtures was not adequate to define the SAR of the hits, a requirement in the evaluation of hits today. Slowly, the industry has turned away from mixture testing and returned to screening single compounds. The high-quality results from single-compound HTS are helping in developing and validating new cheminformatics tools for evaluating hit sets.

The assembly of large screening libraries for HTS requires the development of archival and retrieval systems that can handle solid compounds and compound solutions in microtubes or in plates in either the 96 or the 384 format. Barcodes are used to identify each sample vial, tube, or plate. Databases record which samples are available, what type of sample it is, and how much is available. All solid samples are prepared first as stock (master) solutions (around 5 mg/mL in DMSO) in 2-D bar-coded tubes that are then aliquotted to plates and processed as described earlier. The tubes and plates are usually made of polypropylene for compatibility with DMSO. DMSO is the industry standard solvent for screening libraries because many of the archived compounds are not soluble in water at 5 mg/mL. DMSO, an organic solvent, also has a favorable combination of biological

compatibility and physicochemical properties. The latest efforts with compound libraries center on ways to improve the quality of the solutions starting from the time of their generation through their storage life to their use in an assay. High-quality test compound solutions are critical to the production of high-quality data. In the rush to improve the assays, the handling of compound solutions was a bit neglected for several years, but no longer. A major drawback of the solvent DMSO is its rapid uptake of water (Figure 5.1), a less effective solvent for many of the compounds. If the water content gets too high in the DMSO solutions, some compounds may fall out of solution and not be delivered to the screens, resulting in false negatives. The solutions are stored in microtubes and plates at −5 to –20°C in a dry environment to decrease the uptake of water and to prevent breakdown. If the samples are removed, thawed, and then refrozen, they will take on additional water each time that process is repeated. Each freeze–thaw cycle can add 5% water or more depending on the time of exposure of the DMSO to the air, the relative humidity of the air, the volume of the sample, and the volume of air it is exposed to (Figure 5.1). In the efforts to conserve compound, smaller and smaller aliquots of sample are used for HTS. AT BIPI, the stored aliquots for HTS are 2 μL in a 384-well plate for single use only. Taking the single-use concept up the chain of steps needed to create those final plates requires multiple copies of each step for single use be made and stored. If a company has a million-compound library and makes 10 copies of each of the four stages (master, mother, diluted mother, and daughters), then it will need to store 40 million samples. To store this many samples properly requires the use of automated storage and retrieval systems built by companies like REMP (Oberdiessbach, Switzerland), TAP (Royston, U.K.), RTS International (Frankfurt, Germany), and Matrical (Spokane, WA, U.S.A.). These systems all have large databases associated with them that track the type of vessel, the amount of sample, the location, and other information. The systems that provide this level of flexible storage and retrieval under precisely controlled conditions of temperature and humidity are often costly, but what they protect is almost invaluable.

5.2.3 Biochemical Assays

Biochemical assays are primary tools for the identification of active compounds and for the generation of SARs. As such, they play an indispensable role in the identification of active molecules and their progression to qualified hits, leads, and preclinical candidates. These assays most often employ purified proteins in a defined biochemical system and quantitate binding or catalytic activity in the presence of candidate small-molecule modulators. Early on, HTS practitioners recognized that multistep, workhorse assays of the research laboratories, like the enzyme-linked immunosorbent assay (ELISA), worked poorly or failed outright when transferred to HTS automated systems. In broad terms, the fundamental characteristic that distinguishes a screen (or screening assay) from an ordinary assay is demonstrated compatibility with HTS automation. This compatibility depends on the extent to which the assay is *robust*, that is, one whose output signal is resistant to interference, artifact, and small fluctuations in timing, temperature, and reaction volumes. The most robust screens are composed of few steps using stable reagents, are homogeneous (contain no separation steps, i.e., mix and measure), and produce a stable signal. Because each step of an assay has intrinsic variability of a magnitude dependent on the specific device with which the step is accomplished, fewer steps will decrease overall assay variability and improve robustness. In our experience, the most variable step commonly employed in HTS is microplate washing, almost certainly because of the several variable operations that comprise this seemingly simple step. For example, the efficiency and reproducibility of washing the wells of a microplate can be

affected by the shape of the wells, the rate of wash buffer addition and aspiration, the location of the wash and aspirate nozzles relative to the well surfaces, the composition and physical properties of the wash buffer, and the number and duration of wash cycles. Although contemporary plate washers allow fine adjustment of many of these parameters, the number of combinatorial possibilities makes optimization a difficult and unpredictable process for each distinct liquid class, microplate, and assay. Moreover, achieving identical conditions across an array of 384 delivery and 384 aspiration small-bore nozzles over many cycles represents a significant technical challenge.

Assay variability directly affects the reproducibility of assay results, and reproducibility of results across independent test occasions is a feature expected of screening assays. Some researchers have observed that different assay formats for the same target yield surprisingly different results;[15,16] others have noted good agreement across formats.[17–19] The degree of results concordance within and across different assay formats likely depends upon many factors, including assay design, optimization, implementation, variability, susceptibility to interference, test compound properties, and hit criterion.

An early objective in the development of screening assays, then, was to simultaneously decrease overall variability and increase robustness by reducing the number of steps and eliminating separation steps whenever possible. This need for homogeneous assays lacking wash, filtration, or centrifugation separation steps required the development of breakthrough assay technologies to unleash the full potential of HTS. Perhaps more than any other technological development, except perhaps the automated liquid handler, these miniaturizeable, homogeneous assay technologies have enabled dramatic improvements in HTS efficiency and data quality over the past decade. Detailed descriptions of the manifold HTS-compatible assay technologies in common use have appeared elsewhere;[20–22] a few examples here will suffice to demonstrate the ingenious way in which established biophysical phenomena have been incorporated into robust and sometimes novel assay designs for HTS. The methods favored for biochemical HTS rely in most cases on the detection of a changed state of proximity of two species and associated labels. This proximity change may result from receptor–ligand engagement, may be disrupted by the proteolytic cleavage of a linker peptide, and can be achieved even with molecules not themselves labeled by using tagged antibodies. Homogeneous screening assay formats fall into several categories based on their signal output and whether or not they require a solid phase; several common examples are listed in Table 5.1.

The scintillation proximity assay (SPA),[23,24] an archetype homogeneous HTS technology, was first to be embraced and almost universally adopted by HTS laboratories. Amersham Corporation's debut of SPA in the early 1990s ushered in a new era of homogeneous HTS and provided a benchmark for technologies developed subsequently. SPA technology employs small (2 to 8 μm in diameter) beads whose composition includes a scintillant, so that completion of a molecular binding event that localizes a radiolabeled (most commonly ^{3}H and ^{125}I) ligand near the surface will excite the scintillant and produce photons that can be detected using a luminometer plate reader or imaging device. The first beads were made of either yttrium silicate with cerium ions within the crystal lattice providing scintillant or polyvinyltoluene (PVT) with diphenylanthracine incorporated as scintillant. Beads optimized for lower-volume HTS and CCD imagers are made of yttrium oxide or polystyrene containing a europium red-shifted fluor to minimize quenching.[25] SPA beads are available coated with a variety of different capture molecules, including antibodies, streptavidin, lectins, and metal chelates and can be used to design an extraordinary array of different screening assays, including receptor–ligand, enzyme, and immunoassay. Microplate-based analogs of SPA beads work by the same principle and can be applied to cellular assays.[26,27]

If SPA has shortcomings, and every assay technology does, they include its requirement for pipetting beads, which can be challenging for liquid handlers, and its use of radioisotopes.

TABLE 5.1
Examples of Commercial Homogeneous HTS Technology

Technology Name	Company	Description
AlphaScreen™ (amplified luminescent proximity homogeneous assay)	PerkinElmer Life Sciences	Donor beads produce singlet oxygen that induces fluorescence in proximal acceptor beads; bead–bead proximity results from engagement of binding partners immobilized on the two bead types
ECL (electrochemiluminescence)	Meso Scale Discovery	Binding of ruthenium-labeled moiety to partner on specialized microplate surface produces amplified photon generation
EFC (enzyme fragment complementation)	DiscoveRx	Changed binding state of ligand in conjunction with anti-ligand antibody competition combines two β-galactosidase fragments to make active enzyme that produces signal from fluorogenic or luminogenic substrate
FP (fluorescence polarization)	Several	Interdependent change in molecular size and rotational diffusion of labeled moiety changes FP signal
FRET (fluorescence resonance energy transfer)	Several	Proximity event measured by radiationless energy transfer between donor and acceptor fluorophores followed by photon emission by the latter
HTRF® (homogeneous time-resolved fluorescence); LANCE™ (lanthanide chelate excitation)	Cisbio International; PerkinElmer Life Sciences	Lanthanide TR-FRET using Eu^{3+} cryptate/chelate-donor fluorophore and cross-linked allophycocyanin-acceptor fluorophore
IMAP™ (immobilized metal ion affinity-based fluorescence polarization)	Molecular Devices	Trivalent metal ion containing beads bind to phosphorylated peptide producing FP signal change
IQ® (iron quench)	Pierce Biotechnology, Inc.	Iron compound binds fluorophore-labeled peptides after phosphorylation and quenches fluorescence intensity
Kinase-Glo™; PK-Light™; easylite-Kinase™	Promega; Cambrex; PerkinElmer Life Sciences	Luciferase-based ATP quantitation to measure kinase activity
LanthaScreen™	Invitrogen	Lanthanide TR-FRET using Tb^{3+} chelate-donor fluorophore and fluorescein-acceptor fluorophore
POLARIS™	Chromagen	Reactive biotinylated linker covalently binds phosphorylated peptide, producing FP signal change after avidin binding
QTL Lightspeed™ (quencher-tether-ligand)	QTL Biosystems	Tethered ligand quencher modulates signal from fluorescent polymer in proportion to cleavage of ligand by target enzyme
SPA (scintillation proximity assay)	Amersham Biosciences (GE Healthcare)	Binding of radiolabeled moiety to beads containing scintillant produces luminescent signal
Z′-LYTE™	Invitrogen	Phosphorylation by a kinase of dual labeled FRET peptide changes its susceptibility to proteolytic cleavage

Many scientists would argue that the flexibility and reliability of SPA more than offset these potential disadvantages, and SPA is still in wide use for HTS today. Nevertheless, methods that generate fluorescent and luminescent signals without radioisotopes have become favorites in some HTS laboratories. These fluorescence- and luminescence-based methods,

available in both solution-phase and surface-dependent varieties, have extended the range of homogeneous assays suitable for HTS. For many years, the prospect of fluorescence methods achieving sensitivity comparable to methods using radiolabels in microplate assays was thwarted by the ubiquity of fluorescent substances in the assay environment[28] and by the lack of appropriate microplate readers. Prompt autofluorescence of buffer components, plastic, serum, and test compounds all contribute to a high noise level that can swamp detection limits suggested by quantum-yield calculations for fluorescence-intensity measurement. The application of ratiometric, resonance-energy transfer (RET) and time-resolved (TR) fluorescence methods to biochemical and cellular assays, accompanied by the development of sophisticated multimodal plate readers, have made fluorescence one of the most widely used techniques in HTS today. Certain lanthanides, such as Eu^{3+}, Tb^{3+}, and Sm^{3+}, have fluorescence lifetimes in the microsecond range, and pulsed excitation followed by temporally gated detection of emission avoids most prompt fluorescence background noise, greatly extending the low-end sensitivity of fluorescence. The extension of lanthanide-based time resolution to methods based on fluorescent resonance energy transfer (FRET),[29] exemplified by HTRF®[30,31] and LANCE™,[32] brings the advantages of time-resolution into a homogeneous format that is well suited to HTS.

In some cases, HTS-compatible biochemical assay technologies in the arsenal today are direct descendents of earlier applications of the same technology in diagnostics and biological research. The phenomenon of fluorescence anisotropy, also known as fluorescence polarization (FP), was first applied to assays in the early 1970s and 1980s, and became a mainstay for plate-based formats after the introduction of sensitive plate-based readers by M. Jolley and others.[33–36] FP technology relies on the proximity (binding or dissociation) event producing a significant size change of a fluorophore-labeled moiety, resulting in a measurable change in the polarization of its fluorescence due to a changed rate of rotation in solution. The application of FP to HTS has steadily grown as technological advancements in reader and imager technology, as well as the proliferation of available fluorophores, have bolstered the HTS compatibility of these assays. One limitation of FP for HTS is that the change in polarization of unbound and bound fluorescent probe typically dictates an upper size limit of about 10 kD for the fluorescent species,[35,36] although the use of longer-lived fluorophore probes can push this limit considerably higher.[37] The size of the bound complex can sometimes be increased by the complexation of antibody or avidin moiety to the unlabeled binding partner. The sensitivity of FP in a competitive binding mode is also limited by the K_d of the fluorephore-labeled ligand for the receptor,[38] as is the case for any competitive binding assay.[39] This can be a problem when available probes have K_d values above several hundred nanomolar and significantly lower affinity compounds need to be quantitated. Although FP is a ratiometric method, test compound fluorescence can cause signal interference, significant reduction of which has been accomplished by using longer wavelength fluorophores and by increasing the fluorophore concentration.[40–43] Robust and inexpensive FP screens have been developed using purified component systems as well as lysate and membrane preparations.[44–47] Two FP technologies introduced recently take advantage of metal binding to avoid the necessity for antibodies in the assay of kinases. Immobilized, metal ion, affinity-based fluorescence polarization (IMAP™) was originally developed to assay cyclic nucleic acid phosphodiesterases[48] but has been applied most effectively to kinases.[49] This technology, based on immobilized, metal-affinity chromatography, relies on selective binding of a metal ion-bearing bead to a phosphorylated peptide substrate. Iron quench (IQ®) technology[50] employs an iron-containing compound that binds selectively to phosphoryl groups on fluorescently labeled peptides, reducing overall fluorescence in proportion to the degree of phosphorylation.

Another effective, homogeneous method for the quantitation of kinase activity that requires neither antibodies nor solid phase is the use of a second adenosine triphosphate

(ATP)-dependent enzyme, luciferase, to measure residual ATP following a kinase reaction.[51] This method cannot tolerate ATP concentrations above 10 or 20 μM but has the advantage that the intrinsic ATP hydrolysis activity exhibited by some kinases[52,53] can be measured without the need for a specific polypeptide phosphoryltransfer acceptor. In addition, recent work suggests that screens using this method may identify the most comprehensive set of hits compared with screens using dissociation-enhanced lanthanide fluoroimmunoassay (DELFIA; PerkinElmer Life Sciences) and competitive FP formats.[54] Comparison of kinase assays using luciferase-based ATP measurement and those using a microfluidic method based on electrophoretic or electroosmotic separation showed comparable sensitivity.[18] Another technology that can be used to assay kinases, AlphaScreen™, relies on bead–bead proximity resulting from engagement of binding partners immobilized on the surfaces of separate donor and acceptor beads. Singlet oxygen generated by laser excitation at 680 nm in the donor beads is converted into long-lived luminescent emission at 520 to 620 nm on proximal acceptor beads, resulting in an interference-resistant system whose input light is of a longer wavelength than the output signal.[55] AlphaScreen is a versatile HTS technology that has been used to develop a wide array of screen types.[55–59]

Homogeneous technologies have evolved to be less prone to interferences caused by the presence during signal measurement of test compounds having fluorescent, quenching, or other undesirable properties, yet these technologies remain differentially susceptible to such interferences.[42,60] The most effective way to avoid test-compound interferences is to remove unbound compound before reading the signal, at the cost of losing the benefits unique to homogeneous formats. Among heterogeneous screen formats, the DELFIA is an ELISA analog having a lanthanide fluorophore substituted for the enzyme tag.[61,62] The combination of sensitive TR fluorescence with wash steps renders this format nearly impervious to signal interference. Often generating very large signal windows and sensitivities that rival analogous radiometric assays, DELFIA technology is limited by the ability to effectively coat and wash the wells of the microplate. Although certainly attractive for HTS in 96- and 384-well formats, DELFIAs, in common with other heterogeneous methods requiring wash steps, are not presently as miniaturizeable as homogeneous techniques.

5.2.4 Cellular Assays

Owing to their vastly greater complexity when compared with biochemical assays, cellular assays are often considered an essential bridge to be crossed by a drug candidate on its journey from biochemical testing *in vitro* to evaluation *in vivo*. This view notwithstanding, disagreement about their desirability as front-line HTS can still be heard in some quarters. This disagreement is more about tactics than strategy, because the value of cellular assays run secondary to a biochemical screen is almost universally accepted. Moreover, a recent survey of HTS practices predicts that in 2005 more than half of primary HTS will be cellular.[63] The case made against cellular screens, especially for intracellular targets, is one also used in their favor: Many off-target activities can distract test compounds and confound the identification of selective hits. On the other hand, cellular assays, often called functional assays, are uniquely able to identify orphan-receptor modulators and novel targets affecting signaling pathways of interest. Cellular HTSs are most commonly deployed when the molecular target cannot be isolated or purified in a functional form or when patent coverage of a target prohibits any screen except one run in a whole-cell system against naturally (endogenously) expressed protein. Advantages of cellular screens include the determination of efficacy in the primary screen, the possibility of assessing multiple targets in a pathway, identification of allosteric effectors, and an early indication of cytotoxicity. Challenges for cellular assays include the need to keep DMSO concentrations low, the intensive cell culture required to supply a typical screen, and the possibility for numerous off-target activities of compounds

showing up as hits requiring deconvolution of large numbers of compounds in downstream assays. Common screening targets include cellular receptors[64] and ion channels,[65] and the assay technologies to measure these targets in living cells span the breadth of the cell, including competition binding at the receptor itself, voltage-sensitive dyes, cAMP quantitation, and reporter gene assays measuring transcriptional events. Probably the two most common cellular HTS formats for measuring receptor agonism or antagonism are those measuring receptor antagonism or agonism using Ca^{2+}-sensitive dyes to detect ligand-induced GPCR Ca^{2+} flux, or reporter genes that generate a fluorescent or luminescent signal. Both of these methods are sensitive, straightforward to automate, and compatible with miniaturized formats.[66] Quantitation of cAMP has been accomplished in cellular assays using a variety of homogeneous technologies.[46–48,59,67,68] There are many other cellular parameters that can be measured in assays, including some that use HTS-compatible technology.[64]

5.2.5 The Drive to Miniaturize

The sharp rise in the number of compounds available for screening in the libraries of major pharmaceutical companies caused by combinatorial and automated chemistry efforts has driven both the increase in throughput of HTS and the attendant need for assay formats of lower volume and higher density. With typical pharmaceutical company compound libraries today being in the neighborhood of 10^6 compounds,[63] high-density, low-volume formats are essential to prevent overlong screening campaigns, monopolization of valuable human and automation resources, and excessive consumption of reagents, cells, and test compounds.

A casual observer of HTS evolution during the past decade might conclude that assay wells can never be too small or too closely packed together. There are limits to open systems, however, imposed by the physical properties of liquids and the devices used to manipulate them. The microplate format most widely used for HTS today is 384-well, although 96-, 1536-, and 3456-well formats are used in some settings. It should be pointed out that 384-well format microplates are available as standard volume (30 to 80 μL), and low volume (5 to 30 μL) editions, with working volumes of the latter overlapping with typical 1536-well assay volumes (2 to 10 μL). Once assay volumes reach single-digit microliter levels, several factors impede further scale-down in an open-well format. A major hindrance is evaporation, which can be reduced but not eliminated by using humidified environmental control and innovative plate and well design. Evaporation limits the incubation times for the steps of an assay before loss of water produces a significant impact on reaction component concentrations. Uneven evaporation across an assay microplate can also pose significant problems for extraction of reliable results from all wells in the plate, especially when control wells are affected.[69] A second factor hampering miniaturization below single-digit microliter volumes is the necessity for nanoliter pipetting to assemble these small assay reactions. Nanoliter portions of the full assay reaction will evaporate even faster than microliters, necessitating rapid assay assembly. Meeting these liquid-handling requirements is highly challenging, although nanoliter liquid-handling devices are evolving rapidly.[70,71] Active mixing is more difficult after the jump from 96-to 384-well plates, especially when very different liquid classes are combined. This mixing challenge is exacerbated by the inability to use pipettor trituration at low microliter volumes and may require new approaches such as mixing by application of acoustic energy.[72] Finally, heterogeneous assays are not compatible with low microliter volume assays because of the lack of robust, low-volume wash devices, restricting these low-volume assays to scaleable, homogeneous assay technologies. In spite of the hurdles outlined above, low-volume HTS can be accomplished in certain settings with current technology.[10,11,73,74]

An often-noted paradox in HTS is that as microplate wells and assay volumes have become smaller, automated systems to accommodate them have grown ever larger. This truth raises an important point: Because HTS is not a bottleneck in the drug-discovery

process, decisions about whether to move to lower-volume formats should be informed by careful cost–benefit analysis. Not only is the automated equipment to handle higher-density, lower-volume formats larger, it is substantially more expensive. Savings from reduced consumption of protein reagents, cells, and test compound must be weighed against higher costs for disposables, large initial capital outlay, and possible constraints on compatible assay formats. The HTS world appears to be at a crossroads of scale for reasons outlined above, and it is possible that assembling assays with volumes below several microliters will require abandoning open-well assay plates altogether in favor of closed systems. Such a move could eliminate the above noted inverse relationship between assay volume and HTS infrastructure size. At least two closed systems — microfluidic chip-based systems and well-free plates — are already in use for screening (see Section 5.4).

5.2.6 Data Analysis

Of the three central pillars that support the integrity of the assay endeavor — assay design, assay implementation or data generation, and data analysis or informatics — a great deal of attention is paid to the first two. Although not often the focus of much attention in research laboratories, data handling in the HTS laboratory has become a central concern as the magnitude of HTS output increasingly strained available computational resources. Unlike the convergence of HTS automation technology and processes during the past decade (Section 5.2.1), there are probably as many ways of performing data analysis as there are HTS laboratories, each using a different mix of commercial and in-house developed software, statistical and secondary analyses, and database strategies.[75–77] HTS data-quality monitoring is usually a combination of summary statistics, such as signal/background, signal/noise, and some form of the Z′ statistic,[78] combined with inspection of all data, using powerful visualization tools designed for large data sets. The ultimate goal is to fully automate data processing so that intervention by a scientist is not required except in cases where the data are out of bounds of preestablished metrics. Trends within and across many plates may not be easy to detect using automated methods, however, and interactive visualization is still necessary in many cases. One topic of debate is whether assay-based artifacts, such as pipetting variances, differential evaporation, and the like, should be normalized using control or test wells. Elimination of such artifacts at the assay level is the best option,[79] although this is not always possible. Sophisticated numerical methods to analyze, normalize, and clean large HTS data sets are becoming more common, and reveal clear similarities between HTS and genomics data arrays.[80,81] One particularly interesting method was the application of Fourier transform analysis to 384-well HTS data to detect spatially correlated errors difficult to detect using standard visualization tools.[82]

5.3 The Expanding Use of Higher Throughput Assays

The components that have made HTS a smooth-running machine combined with the high-quality data that has been shown to support SAR have led to natural expansions of the HTS arena. Originally, an HTS operation simply identified active compounds in a mixture that was tested at a single concentration. The mixture was deconvoluted into its various components, and each component was tested for activity. This was considered a confirmed active. Most companies now test single compounds, so a second testing of the original solution is considered to be the confirmation. As the collections grew dramatically in size from 100,000 to 200,000 compounds to more than a million compounds, the downstream

processes, including HTL, were flooded by thousands of active hits requiring further characterization. All HTS data sets have both false positives and false negatives. The first test to eliminate false positives in the hit set is to test the compounds for dose responsiveness. This testing step was reassigned from the therapeutic area to HTS when the hit sets went from handfuls of hits to thousands of hits being available at one time. The test for dose responsiveness progressed to the determination of the concentration that inhibits 50% of the response (IC_{50}) values as the quality of screening assay data rose. This usually meant going from five-point concentration curves to, typically, ten points. Concentration curves can be made by diluting a stock concentration out from well to well across a 96-well plate with an eight-tip pipettor. A fast and efficient way to prepare dose–response samples for an HTS operation is to leverage the 96- or 384-tip pipettors and to make dose–response plates by diluting from plate to plate instead of from well to well within each plate (Figure 5.2). This method reduces the exposure of the DMSO solutions to the wet-air environment, and control wells on each plate are used to normalize for any DMSO concentration differences and plate-to-plate variability. Different groups will put controls either on both end columns or the first and last columns or the middle and last columns as they create 384-well plates from four of the 96-well plates that usually have only one column of controls. This plate-to-plate concept for compound dilutions has been fully validated against the well-to-well dilution scheme at BIPI. Today, the initial stock solutions are diluted in neat DMSO and plated out as the appropriate 2 µL aliquots to be diluted into aqueous buffer immediately before testing. This procedure helps to keep the compounds in solution until they are diluted into the aqueous buffer and ensures that even if the highest concentrations come out of solution, the lower concentrations will have the best chance of being correct. Finally, compounds rarely exhibit sticking to pipettor tips and carryover when dilutions are done in DMSO, whereas

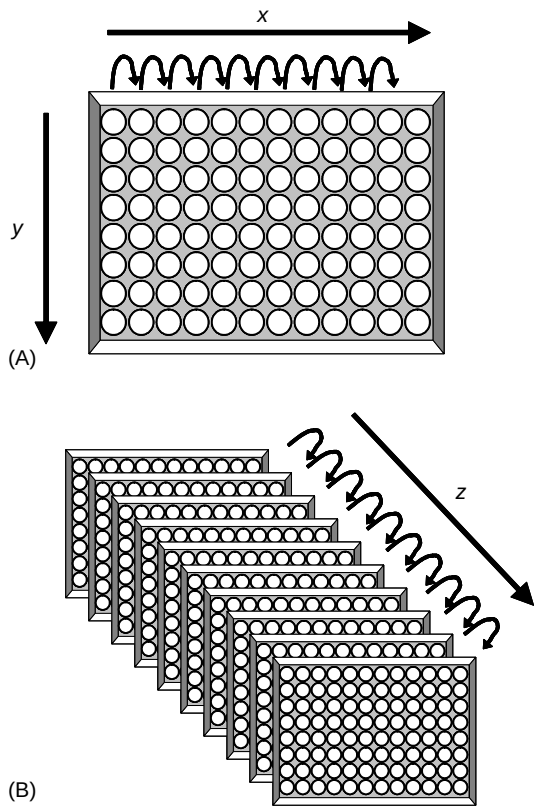


FIGURE 5.2
Leveraging 96- or 384-tip array pipettors for efficient compound serial dilution. (A) typical intra-plate (*x*–*y*), serial dilutions accomplished using an 8- or 12-tip pipettor; (B) more efficient interplate (*z*), serial dilutions using an array pipettor.

this can be a serious problem for certain compounds diluted in assay buffer and can cause aberrant and misleading IC_{50} curves (Figure 5.3).

The project teams or therapeutic area teams using the IC_{50} values generated by the HTS assays often requested the HTS group to test all hits in a selectivity assay or some other assay that would qualify the hit set. Ideally, this counterscreen would also eliminate false positives caused by an assay artifact that, in some cases, can produce an authentic-looking dose–response curve. It was easier for HTS to test all of the hits at the same time rather than just the dose–responsive hits — the procedure usually followed by therapeutic-area laboratories. The testing of large hit sets in one assay on the same day under identical conditions will produce less error in the overall data because there are no additional errors such as those that result from day-to-day variation. In areas where there are closely related assays, the HTS group might perform several counterscreens to reduce the hit set, which may number to thousands of compounds, low hundreds, or even lower. After this qualification of the entire hit set, the HTL or therapeutic-area chemists evaluate all of the data to determine which compounds are of further interest. This step might include clustering of the hits (groups of like compounds based on their chemical structures and properties) to eliminate compounds that fail certain physical-property rules. This is also the first time the compounds are evaluated for SAR patterns. There could be hundreds of interesting compounds that have not been tested as part of the HTS library, and the chemist might want to confirm SAR patterns and check for potential false negatives. This next round of testing is performed by HTS to get all of these compounds tested in full dose–response curves under identical conditions again in both the original screen and any counterscreens. At BIPI, the HTS group now supports the primary assay and its counterscreens through the HTL process when the hit sets are sufficiently large. Today, automated chemistry is often used to expand hit sets if the hits are not

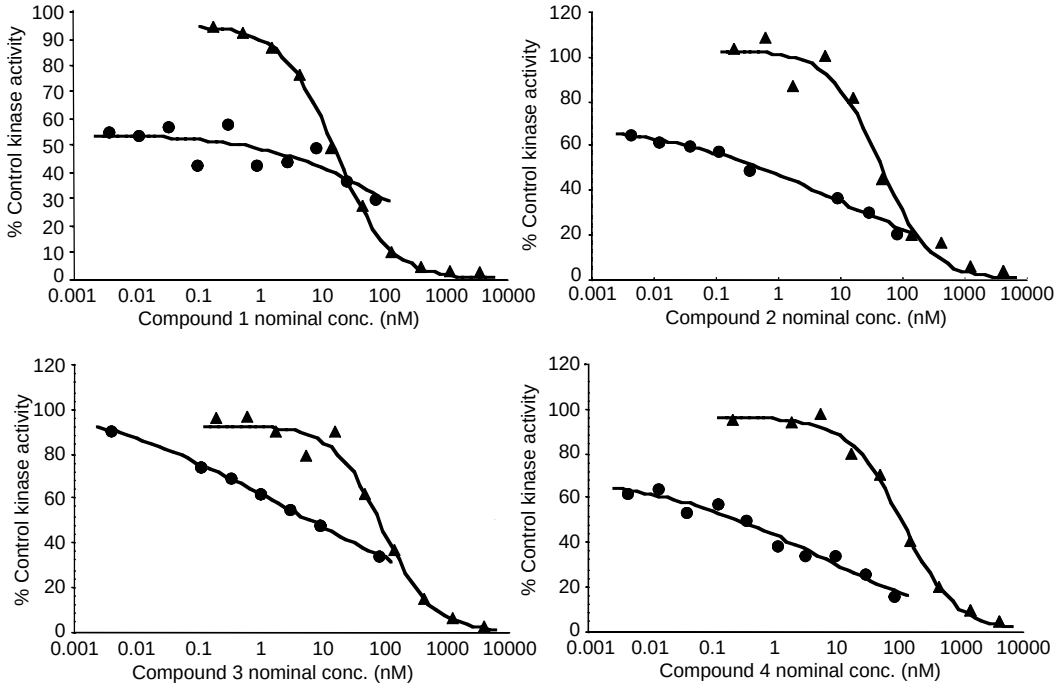


FIGURE 5.3
Comparison of inhibition curves generated by four compounds following serial dilution in assay buffer containing 1% (v/v) DMSO (circles) and in neat DMSO (triangles) in a protein tyrosine kinase assay.

well represented in the chemical library to help define the SAR and to validate that the chemistry process is reproducible. SAR and a tractable chemical synthesis are two of the chemical criteria needed to move a series of chemistry from the hit stage to the lead stage in drug discovery. The total set of requisites for a lead series will vary from company to company but includes both the biology (potency and selectivity) and the chemistry (SAR and tractable synthesis) criteria mentioned previously. Automated assays performed in extended HTS groups (the Biomolecular Screening Group [BMSG] at BIPI, the Lead Identification Group [LIG] at some companies) help the HTL process reach decision points much faster. Today, the typical HTL phase of drug discovery is less than a year as compared with the 18 to 24 months it took a few years ago. Additional lead criteria often include results from *in vitro* assays that evaluate the potential for drug–drug interactions (cyp450 assays), the drug metabolism rates (liver microsome assays), the permeability and active transport rates (Caco-2 cell assays), and other absorption, distribution, metabolism and elimination (ADME) properties. At BIPI, we call this set of assays *target-independent profiling* to distinguish them from all of the previous testing directed at the specific target. These *in vitro* assays have been automated to handle larger volumes of compounds and are often called high throughput. Although, they do not process nearly the number of compounds as does the primary screen, the automated assays do run at a higher level of throughput and, as such, are considered HTS within their context. The BMSGs or LIGs may also perform these assays because they are the most familiar with the automation levels and can apply the best level of automation for the needed throughputs. Performing the assays in this manner, as mentioned before, eliminates many of the variability factors and allows for direct comparison of the data from compound to compound as well as from series to series, which is an important feature for the lead series designation. The drug-discovery process must ensure that drug-like features be incorporated into the compound lead series as early as possible. It is not enough to produce potent compounds to produce drugs. It is a delicate balance between potency, selectivity, and ADME features that leads to the best drugs. The high-quality data generated by HTS assays today has demonstrated that many of the drug-discovery processes can benefit from the use of the HTS techniques past the initial, primary HTS screen. These processes can even be used in the LO phase of drug discovery. Most of the *in vitro* assays developed for the target, as well as the ADME assays mentioned for obtaining lead status, continue to support the drug-discovery program through the LO phase. This process provides a continuum from HTS to HTL to LO, which significantly decreases timelines from the past when assays were handed off from group to group, each of which often developed significantly different procedures requiring months to develop and validate. The total effect of incorporating HTS procedures into the drug-discovery process will be not only a reduction in the timelines of drug discovery but the ability to produce more data to drive critical decisions about compound progression. In the longer term, this overall process change should provide better compounds to the development phase of the drug-discovery process. The overall drug-discovery process has traditionally taken 15 years or more. The incorporation of HTS techniques into this process is a relatively new development that is still ongoing. Several more years will be required to evaluate its full impact on the discovery of new drugs.

5.4 The Road Ahead

Future developments in any area of technology are notoriously difficult to foretell, although they often proceed incrementally via logical extensions of current technology. It is safe to say that in the HTS arena there will continue to be incremental advancement in

throughput, miniaturization, sensitivity, and robustness. The use of fluorescence in HTS will likely continue to expand, as new techniques like modulation and polarization sensing begin to be applied to microplate assay systems.[83] Single-molecule detection systems suitable for HTS should be near at hand as well, because operational systems are in place.[84–88] In cellular assays, single-cell electrophysiological measurements for ion-channel screens are at the doorstep of screen-compatible throughputs,[89] and high-content screening using confocal imaging to measure complex cellular phenomenon[90,91] will certainly expand beyond its current niche status in HTS.

As discussed in Section 5.2.5, HTS has reached a crossroad at the 2 to 10 μL assay volume. Will this be the lower limit of microplate-based formats, or will evaporation in nanoliter reaction wells be controlled using assay systems housed in hyperbaric chambers? Two recently developed formats have moved to submicroliter assay volumes by abandoning the microplate well and moving to microfluidic or well-free surfaces. Elegant, closed assay systems using microfluidic chips[18,92] and well-free sheets[93,94] have been used for screening, but they are not yet versatile enough to be able to handle every class of target that an HTS group might see in a year of screening. In both cases, a primary difficulty lies in the delivery of large compound collections, mostly stored in 96- and 384-well formats, to the nanoworld of these applications without compound waste and cross-contamination. In addition, methods that rely on compounds dried from DMSO stocks redissolving in aqueous buffer to enter the assay matrix may not deliver reliable concentrations across diverse sets of test compounds with widely different solubility properties. When technologies are developed to better bridge this compound storage and assay delivery interface, these closed systems will be more compelling for general deployment in the HTS laboratory.

Assay development of tomorrow will likely be more and more subject to enhancement through automation, as has already begun to happen with the application of design of experiments (DOE) to assay development.[95] Automated assay optimization (AAO) combines the power of automated pipetting with sophisticated combinatorial design and statistical analysis to test many more components and their effect on each other than is possible using manual methods. One microplate-based technology on the threshold of HTS readiness that has the potential to radically shrink assay development times is label-free detection of small-molecule binding to immobilized macromolecular targets using waveguide sensor technology.[96,97] We may also look ahead for the rise of virtual screening performed *in silico*, accompanied by real screening of small subsets of compounds, especially as structures become available from small-scale protein production and automation-assisted crystallization. In the more distant future, when the protein-folding problem is solved and the tertiary structure and conformational states of proteins can be deduced accurately from primary sequence alone, virtual screening will have gained its ascendancy and wet-lab screening may become a vestigial remnant of present-day screening factories.

5.5 Conclusions

High-throughput screening today is a well-organized process used by major pharmaceutical companies in their quest to discover new drugs. Although some have questioned the value of the investments made in new technologies for drug discovery,[98] the 74 compounds in the clinic that can trace their roots back to an HTS effort[63] provide ample justification for these investments. Potential drugs continue to fail in the clinic for various reasons, but these failures have nothing to do with their start from HTS. The one constant

in HTS over the past 10 to 15 years has been change. HTS has made needed improvements in the process to raise the quality of the data produced. The decision to improve the quality of the HTS data was perhaps the biggest and most dramatic decision made during these years. This decision redefined HTS and moved it away from the initial concept of a crude screen to see what fell out. Today, the product of HTS is a well-defined hit set that can be easily validated and rapidly moved forward to at least one, and preferably more than one, lead series. It has taken time to work out all the logistical and philosophical issues with the HTS process because the process was in full production at the same time as it needed changing. This was like trying to change a flat tire on a car while it is still moving. All the skills and tools that have been brought together for today's HTS will continue to evolve and improve to make the process even better. In the future, science will provide a greater degree of knowledge about the biological targets. This knowledge may allow for more directed or virtual screening instead of whole-collection screening for some targets. Today, the structural knowledge about the protein target required for virtual screening is usually not available when the decision to run the HTS is made. However, many companies are presently conducting experiments to see which types of targets may be applicable for subset screening. Several companies are performing gene-family screening, particularly for kinases, with selected kinase inhibitor-like libraries. Drug discovery is the screening of samples against targets, and the automated assays and infrastructure built for HTS will continue to support this screening. The numbers of compounds or the throughput rate may change, but the HTS process is flexible and can screen whatever is needed by the target, from the full library to selected subsets, to find those hits that lead off the drug-discovery process. But, as pointed out in the introduction, HTS was never defined by the screening of specific numbers of compounds. It is, and will continue to be, a means by which biological targets are evaluated for their "drugability" in the most efficient way possible using the latest assay technologies and appropriate automation. So whatever HTS might be in the future, it will always be drug discovery.

References

1. Devlin, J.P., Carter, C., and Homon, C.A., High capacity screening: natural products as immune regulators, in *Economic and Medicinal Plant Research*, Wagner, H. and Farnsworth, N.R., Eds., Academic Press, San Diego, 1991, chap. 4.
2. Fernandes, P.B., Moving into the third millennium after a century of screening, in *Handbook of Drug Screening*, Seethala, R. and Fernandes, P.B., Eds., Marcel Dekker, New York, 2001, chap. 1.
3. Hargrave, K.D., Proudfoot, J.R., Grozinger, K.G., Cullen, E., Kapadia, S.R., Patel, U.R., Fuchs, V.U., Mauldin, S.C., Vitous, J., and Behnke, M.L., Novel non-nucleoside inhibitors of HIV-1 reverse transcriptase. 1. Tricyclic pyridobenzo- and dipyridodiazepinones, *J. Med. Chem.*, 34, 2231, 1991.
4. Farina, P.R., Homon, C.A., Lazer, E.S., and Parks, T.P., Discovery of BIRM 270: a new class of leukotriene biosynthesis inhibitors, in *The Search for Anti-Inflammatory Drugs, Case Histories from Concept to Clinic,* Merluzzi, V.J. and Adams, J., Eds., Birkhèauser, Boston, 1995, chap. 9.
5. Regan, J., Breitfelder, S., Cirillo, P., Gilmore, T., Graham, A.G., Hickey, E., Klaus, B., Madwed, J., Moriak, M., Moss, N., Pargellis, C., Pav, S., Proto, A., Swinamer, A., Tong, L., and Torcellini, C., Pyrazole urea-based inhibitors of p38 MAP kinase: from lead compound to clinical candidate, *J. Med. Chem.*, 45, 2994, 2002.
6. Mere, L., Bennett, T., Coassin, P., England, P., Hamman, B., Rink, T., Zimmerman, S., and Negulescu, P., Miniaturized FRET assays and microfluidics: key components for ultra-high-throughput screening, *Drug Discov. Today*, 4, 363, 1999.

7. Minor, L.K., Westover, L., Kong, Y., Patel, D., Wildey, M.J., Conway, B.R., and Demarest, K.T., Validation of a cell-based screen for insulin receptor modulators by quantification of insulin receptor phosphorylation, *J. Biomol. Screen.*, 8, 439, 2003.
8. Xin, H., Bernal, A., Amato, F.A., Pinhasov, A., Kauffman, J., Brenneman, D.E., Derian, C.K., ndrade-Gordon, P., Plata-Salaman, C.R., and Ilyin, S.E., High-throughput siRNA-based functional target validation, *J. Biomol. Screen.*, 9, 286, 2004.
9. Huang, K.S., Holmgren, J., Reik, L., Lucas-McGady, D., Roberts, J., Liu, C.M., and Levin, W., High-throughput methods for measuring heparanase activity and screening potential antimetastatic and anti-inflammatory agents, *Anal. Biochem.*, 333, 389, 2004.
10. Peekhaus, N.T., Ferrer, M., Chang, T., Kornienko, O., Schneeweis, J.E., Smith, T.S., Hoffman, I., Mitnaul, L.J., Chin, J., Fischer, P.A., Blizzard, T.A., Birzin, E.T., Chan, W., Inglese, J., Strulovici, B., Rohrer, S.P., and Schaeffer, J.M., A beta-lactamase-dependent Gal4-estrogen receptor beta transactivation assay for the ultra-high throughput screening of estrogen receptor beta agonists in a 3456-well format, *Assay Drug Dev. Technol.*, 1, 789, 2003.
11. Kornienko, O., Lacson, R., Kunapuli, P., Schneeweis, J., Hoffman, I., Smith, T., Alberts, M., Inglese, J., and Strulovici, B., Miniaturization of whole live cell-based GPCR assays using microdispensing and detection systems, *J. Biomol. Screen.*, 9, 186, 2004.
12. Bindseil, K.U., Jakupovic, J., Wolf, D., Lavayre, J., Leboul, J., and van der Pyl, D., Pure compound libraries; a new perspective for natural product based drug discovery, *Drug Discov. Today*, 6, 840, 2001.
13. Lipinski, C.A., Drug-like properties and the causes of poor solubility and poor permeability, *J. Pharmacol. Toxicol. Methods*, 44, 235, 2000.
14. Devlin, J.J., Liang, A., Trinh, L., Polokoff, M.A., Senator, D., Zheng, W., Kondracki, J., Kretschmer, P.J., Morser, J., Lipson, S.E., Spann, R., Loughlin, J.A., Dunn, K.V., and Morrissey, M.M., High capacity screening of pooled compounds: identification of active compound without re-assay of pool members, *Drug Dev. Res.*, 37, 80, 1996.
15. Sills, M.A., Weiss, D., Pham, Q., Schweitzer, R., Wu, X., and Wu, J.J., Comparison of assay technologies for a tyrosine kinase assay generates different results in high throughput screening, *J. Biomol. Screen.*, 7, 191, 2002.
16. Wu, X., Glickman, J.F., Bowen, B.R., and Sills, M.A., Comparison of assay technologies for a nuclear receptor assay screen reveals differences in the sets of identified functional antagonists, *J. Biomol. Screen.*, 8, 381, 2003.
17. Hubert, C.L., Sherling, S.E., Johnston, P.A., and Stancato, L.F., Data concordance from a comparison between filter binding and fluorescence polarization assay formats for identification of ROCK-II inhibitors, *J. Biomol. Screen.*, 8, 399, 2003.
18. Pommereau, A., Pap, E., and Kannt, A., Two simple and generic antibody-independent kinase assays: comparison of a bioluminescent and a microfluidic assay format, *J. Biomol. Screen.*, 9, 409, 2004.
19. Newman, M. and Josiah, S., Utilization of fluorescence polarization and time resolved fluorescence resonance energy transfer assay formats for SAR studies: SRC kinase as a model system, *J. Biomol. Screen.*, 9, 525, 2004.
20. Devlin, J.P., Ed., *High Throughput Screening, the Discovery of Bioactive Substances*, Marcel Dekker, New York, 1997.
21. Janzen, W.P., Ed., *High Throughput Screening, Methods and Protocols*, Humana Press, NJ, 2002.
22. Seethala, R. and Fernandes, P.B., Eds., *Handbook of Drug Screening*, Marcel Dekker, New York, 2001.
23. Picardo, M. and Hughes, K.T., Scintillation proximity assays, in *High Throughput Screening, The Discovery of Bioactive Substances*, Devlin, J.P., Ed., Marcel Dekker, New York, 1997, chap. 16.
24. Cook, N.D., Scintillation proximity assay: a versatile high-throughput screening technology, *Drug Discov. Today*, 1, 287, 2004.
25. Zheng, W., Carroll, S.S., Inglese, J., Graves, R., Howells, L., and Strulovici, B., Miniaturization of a hepatitis C virus RNA polymerase assay using a −102°C cooled CCD camera-based imaging system, *Anal. Biochem.*, 290, 214, 2001.
26. Harris, D.W., Kenrick, M.K., Pither, R.J., Anson, J.G., and Jones, D.A., Development of a high-volume in situ mRNA hybridization assay for the quantification of gene expression utilizing scintillating microplates, *Anal. Biochem.*, 243, 249, 1996.

27. Brown, B.A., Cain, M., Broadbent, J., Tompkins, S., Henrich, G., Joseph, R., Casto, S., Harney, H., Greene, R., Delmendo, R., and Ng, S., FlashPlate™ Technology, in *High Throughput Screening*, *The Discovery of Bioactive Substances*, Devlin, J.P., Ed., Marcel Dekker, New York, 1997, chap. 17.
28. Hemmilä, I.A., Fluorescence immunoassays, in *Applications of Fluorescence in Immunoassays*, Winefordner, J.D. and Kolthoff, I.M., Eds., Wiley, New York, 1991, chap. 8.
29. Selvin, P.R., Principles and biophysical applications of lanthanide-based probes, *Annu. Rev. Biophys. Biomol. Struct.*, 31, 275, 2002.
30. Kolb, A.J., Burke, J.W., and Mathis, G., Homogeneous, time-resolved fluorescence method for drug discovery, in *High Throughput Screening, The Discovery of Bioactive Substances*, Devlin, J.P., Ed., Marcel Dekker, New York, 1997, chap. 19.
31. Mathis, G., HTRF® Technology, *J. Biomol. Screen.*, 4, 309, 1999.
32. Hemmilä, I., LANCE™: homogeneous assay platform for HTS, *J. Biomol. Screen.*, 4, 303, 1999.
33. Jolley, M.E., Fluorescence polarization immunoassay for the determination of therapeutic drug levels in human plasma, *J. Anal. Toxicol.*, 5, 236, 1981.
34. Nasir, M.S. and Jolley, M.E., Fluorescence polarization: an analytical tool for immunoassay and drug discovery, *Comb. Chem. High Throughput Screen.*, 2, 177, 1999.
35. Owicki, J.C., Fluorescence polarization and anisotropy in high throughput screening: perspectives and primer, *J. Biomol. Screen.*, 5, 297, 2000.
36. Jolley, M.E., Fluorescence polarization assays for the detection of proteases and their inhibitors, *J. Biomol. Screen.*, 1, 33, 1996.
37. Guo, X.Q., Castellano, F.N., Li, L., and Lakowicz, J.R., Use of a long-lifetime Re(I) complex in fluorescence polarization immunoassays of high-molecular-weight analytes, *Anal. Chem.*, 70, 632, 1998.
38. Huang, X., Fluorescence polarization competition assay: the range of resolvable inhibitor potency is limited by the affinity of the fluorescent ligand, *J. Biomol. Screen.*, 8, 34, 2003.
39. Ekins, R.P., Radioimmunoassay and saturation analysis. Basic principles and theory, *Br. Med. Bull.*, 30, 3, 1974.
40. Banks, P., Gosselin, M., and Prystay, L., Impact of a red-shifted dye label for high throughput fluorescence polarization assays of G protein-coupled receptors, *J. Biomol. Screen.*, 5, 329, 2000.
41. Turconi, S., Shea, K., Ashman, S., Fantom, K., Earnshaw, D.L., Bingham, R.P., Haupts, U.M., Brown, M.J., and Pope, A.J., Real experiences of uHTS: a prototypic 1536-well fluorescence anisotropy-based uHTS screen and application of well-level quality control procedures, *J. Biomol. Screen.*, 6, 275, 2001.
42. Turek-Etienne, T.C., Small, E.C., Soh, S.C., Xin, T.A., Gaitonde, P.V., Barrabee, E.B., Hart, R.F., and Bryant, R.W., Evaluation of fluorescent compound interference in 4 fluorescence polarization assays: 2 kinases, 1 protease, and 1 phosphatase, *J. Biomol. Screen.*, 8, 176, 2003.
43. Turek-Etienne, T.C., Lei, M., Terracciano, J.S., Langsdorf, E.F., Bryant, R.W., Hart, R.F., and Horan, A.C., Use of red-shifted dyes in a fluorescence polarization AKT kinase assay for detection of biological activity in natural product extracts, *J. Biomol. Screen.*, 9, 52, 2004.
44. Banks, P., Gosselin, M., and Prystay, L., Fluorescence polarization assays for high throughput screening of G protein-coupled receptors, *J. Biomol. Screen.*, 5, 159, 2000.
45. Harris, A., Cox, S., Burns, D., and Norey, C., Miniaturization of fluorescence polarization receptor-binding assays using CyDye-labeled ligands, *J. Biomol. Screen.*, 8, 410, 2003.
46. Prystay, L., Gagne, A., Kasila, P., Yeh, L.A., and Banks, P., Homogeneous cell-based fluorescence polarization assay for the direct detection of cAMP, *J. Biomol. Screen.*, 6, 75, 2001.
47. Allen, M., Hall, D., Collins, B., and Moore, K., A homogeneous high throughput nonradioactive method for measurement of functional activity of Gs-coupled receptors in membranes, *J. Biomol. Screen.*, 7, 35, 2002.
48. Huang, W., Zhang, Y., and Sportsman, J.R., A fluorescence polarization assay for cyclic nucleotide phosphodiesterases, *J. Biomol. Screen.*, 7, 215, 2002.
49. Gaudet, E.A., Huang, K.S., Zhang, Y., Huang, W., Mark, D., and Sportsman, J.R., A homogeneous fluorescence polarization assay adaptable for a range of protein serine/threonine and tyrosine kinases, *J. Biomol. Screen.*, 8, 164, 2003.

50. McCauley, T.J., Mathrubutham, M., Morgan, A.G., Onken, J., Stanaitis, M.L., and Millis, S.Z., IQ® Technology: an automated, high-throughput screening assay for kinase, phosphatase, and protease targets, *J. Assoc. Lab. Automation*, 9, 171, 2004.
51. Lundin, A., Use of firefly luciferase in ATP-related assays of biomass, enzymes, and metabolites, *Methods Enzymol.*, 305, 346, 2000.
52. Knight, W.B. and Cleland, W.W., Thiol and amino analogues as alternate substrates for glycerokinase from *Candida mycoderma*, *Biochemistry*, 28, 5728, 1989.
53. Chen, G., Porter, M.D., Bristol, J.R., Fitzgibbon, M.J., and Pazhanisamy, S., Kinetic mechanism of the p38-alpha MAP kinase: phosphoryl transfer to synthetic peptides, *Biochemistry*, 39, 2079, 2000.
54. Kashem, M., Yingling, J., Nelson, R., and Homon, C., New tools for screening protein kinases: a comparative study, presented at *9th Annual Conference and Exhibition of the Society for Biomolecular Screening*, Portland, OR, September 21–25, 2003.
55. Rouleau, N., Turcotte, S., Mondou, M.H., Roby, P., and Bosse, R., Development of a versatile platform for nuclear receptor screening using AlphaScreen, *J. Biomol. Screen.*, 8, 191, 2003.
56. Wilson, J., Rossi, C.P., Carboni, S., Fremaux, C., Perrin, D., Soto, C., Kosco-Vilbois, M., and Scheer, A., A homogeneous 384-well high-throughput binding assay for a TNF receptor using alphascreen technology, *J. Biomol. Screen.*, 8, 522, 2003.
57. Peppard, J., Glickman, F., He, Y., Hu, S.I., Doughty, J., and Goldberg, R., Development of a high-throughput screening assay for inhibitors of aggrecan cleavage using luminescent oxygen channeling (AlphaScreen), *J. Biomol. Screen.*, 8, 149, 2003.
58. Von, L.A., Kumpf, R., Menzel, S., Reulle, D., Griebel, R., Valler, M.J., and Büttner, F.H., Miniaturization and validation of a high-throughput serine kinase assay using the alpha screen platform, *J. Biomol. Screen.*, 9, 719, 2004.
59. Gabriel, D., Vernier, M., Pfeifer, M.J., Dasen, B., Tenaillon, L., and Bouhelal, R., High throughput screening technologies for direct cyclic AMP measurement, *Assay Drug Dev. Technol.*, 1, 291, 2003.
60. Comley, J., Assay interference: a limiting factor in HTS?, *Drug Discov. World*, 4, 91, 2003.
61. Hemmilä, I., Dakubu, S., Mukkala, V.M., Siitari, H., and Lövgren, T., Europium as a label in time-resolved immunofluorometric assays, *Anal. Biochem.*, 137, 335, 1984.
62. Braunwalder, A.F., Yarwood, D.R., Sills, M.A., and Lipson, K.E., Measurement of the protein tyrosine kinase activity of c-SRC using time-resolved fluorometry of europium chelates, *Anal. Biochem.*, 238, 159, 1996.
63. Fox, S., Farr-Jones, S., Sopchak, L., Boggs, A., and Comley, J., High-throughput screening: searching for higher productivity, *J. Biomol. Screen.*, 9, 354, 2004.
64. Seethala, R., Receptor screens for small molecule agonist and antagonist discovery, in *Handbook of Drug Screening*, Seethala, R. and Fernandes, P.B., Eds., Marcel Dekker, New York, 2001, chap. 7.
65. Falconer, M., Smith, F., Surah-Narwal, S., Congrave, G., Liu, Z., Hayter, P., Ciaramella, G., Keighley, W., Haddock, P., Waldron, G., and Sewing, A., High-throughput screening for ion channel modulators, *J. Biomol. Screen.*, 7, 460, 2002.
66. Johnston, P.A. and Johnston, P.A., Cellular platforms for HTS: three case studies, *Drug Discov. Today*, 7, 353, 2002.
67. Weber, M., Ferrer, M., Zheng, W., Inglese, J., Strulovici, B., and Kunapuli, P., A 1536-well cAMP assay for Gs- and Gi-coupled receptors using enzyme fragmentation complementation, *Assay Drug Dev. Technol.*, 2, 39, 2004.
68. Golla, R. and Seethala, R., A homogeneous enzyme fragment complementation cyclic AMP screen for GPCR agonists, *J. Biomol. Screen.*, 7, 515, 2002.
69. Oldenburg, K.R., Kariv, I., Zhang, J., Chung, T.D.Y., and Lin, S., Assay miniaturization: developing technologies and assay formats, in *Handbook of Drug Screening*, Seethala, R. and Fernandes, P.B., Eds., Marcel Dekker, New York, 2001.
70. Comley, J., Continued miniaturisation of assay technologies drives market for nanolitre dispensing, *Drug Discov. Today*, 5, 43, 2004.
71. Hsieh, H.B., Fitch, J., White, D., Torres, F., Roy, J., Matusiak, R., Krivacic, B., Kowalski, B., Bruce, R., and Elrod, S., Ultra-high-throughput microarray generation and liquid dispensing using multiple disposable piezoelectric ejectors, *J. Biomol. Screen.*, 9, 85, 2004.

72. Hobbs, S., Patel, P., Coyne, C., and Oldenburg, K., High-throughput, low-volume determination of thermodynamic solubility using sonication and micro parallel liquid chromatography, presented at 9th *Annual Conference and Exhibition of the Society for Biomolecular Screening*, Portland, OR, September 21–25, 2003.
73. Wright, P.A., Boyd, H.F., Bethell, R.C., Busch, M., Gribbon, P., Kraemer, J., Lopez-Calle, E., Mander, T.H., Winkler, D., and Benson, N., Development of a 1-μL scale assay for mitogen-activated kinase kinase 7 using 2-D fluorescence intensity distribution analysis anisotropy, *J. Biomol. Screen.*, 7, 419, 2002.
74. Von, L.A., Kumpf, R., Menzel, S., Reulle, D., Griebel, R., Valler, M.J., and Büttner, F. H., Miniaturization and validation of a high-throughput serine kinase assay using the alpha screen platform, *J. Biomol. Screen.*, 9, 719, 2004
75. Roberts, B.R., Screening informatics: adding value with meta-data structures and visualization tools, *Drug Discov. Today*, 1, 10, 2004.
76. Gunter, B., Brideau, C., Pikounis, B., and Liaw, A., Statistical and graphical methods for quality control determination of high-throughput screening data, *J. Biomol. Screen.*, 8, 624, 2003.
77. Brideau, C., Gunter, B., Pikounis, B., and Liaw, A., Improved statistical methods for hit selection in high-throughput screening, *J. Biomol. Screen.*, 8, 634, 2003.
78. Zhang, J.-H., Chung, T.D.Y., and Oldenburg, K.R., A simple statistical parameter for use in evaluation and validation of high throughput screening assays, *J. Biomol. Screen.*, 4, 67, 1999.
79. Lundholt, B.K., Scudder, K.M., and Pagliaro, L., A simple technique for reducing edge effect in cell-based assays, *J. Biomol. Screen.*, 8, 566, 2003.
80. Glick, M., Klon, A.E., Acklin, P., and Davies, J.W., Enrichment of extremely noisy high-throughput screening data using a naive Bayes classifier, *J. Biomol. Screen.*, 9, 32, 2004.
81. Nadon, R. and Shoemaker, J., Statistical issues with microarrays: processing and analysis, *Trends Genet.*, 18, 265, 2002.
82. Root, D.E., Kelley, B.P., and Stockwell, B.R., Detecting spatial patterns in biological array experiments, *J. Biomol. Screen.*, 8, 393, 2003.
83. Lakowicz, J.R., Gryczynski, I., and Gryczynski, Z., Novel fluorescence sensing methods for high throughput screening, *J. Biomol. Screen.*, 5, 123, 2000.
84. Moore, K.J., Turconi, S., Ashman, S., Ruediger, M., Haupts, U., Emerick, V.V, and Pope, A.J., Single molecule detection technologies in miniaturized high throughput screening: fluorescence correlation spectroscopy, *J. Biomol. Screen.*, 4, 335, 1999.
85. Eggeling, C., Brand, L., Ullmann, D., and Jäger, S., Highly sensitive fluorescence detection technology currently available for HTS, *Drug Discov. Today*, 8, 632, 2003.
86. Jäger, S., Garbow, N., Kirsch, A., Preckel, H., Gandenberger, F.U., Herrenknecht, K., Rudiger, M., Hutchinson, J.P., Bingham, R.P., Ramon, F., Bardera, A., and Martin, J., A modular, fully integrated ultra-high-throughput screening system based on confocal fluorescence analysis techniques, *J. Biomol. Screen.*, 8, 648, 2003.
87. Haupts, U., Rudiger, M., Ashman, S., Turconi, S., Bingham, R., Wharton, C., Hutchinson, J., Carey, C., Moore, K.J., and Pope, A.J., Single-molecule detection technologies in miniaturized high-throughput screening: fluorescence intensity distribution analysis, *J. Biomol. Screen.*, 8, 19, 2003.
88. Skinner, G.M. and Visscher, K., Single-molecule techniques for drug discovery, *Assay Drug Devel. Technol.*, 2, 397, 2004.
89. Wood, C., Williams, C., and Waldron, G.J., Patch clamping by numbers, *Drug Discov. Today*, 9, 434, 2004.
90. Ramm, P., Alexandrov, Y., Cholewinski, A., Cybuch, Y., Nadon, R., and Soltys, B.J., Automated screening of neurite outgrowth, *J. Biomol. Screen.*, 8, 7, 2003.
91. Mastyugin, V., McWhinnie, E., Labow, M., and Buxton, F., A quantitative high-throughput endothelial cell migration assay, *J. Biomol. Screen.*, 9, 712, 2004.
92. Sundberg, S.A., Chow, A., Nikiforov, T., and Wada, H.G., Microchip-based systems for target validation and HTS, *Drug Discov. Today*, 5, 92, 2000.
93. Anderson, S.N., Cool, B.L., Kifle, L., Chiou, W., Egan, D.A., Barrett, L.W., Richardson, P.L., Frevert, E.U., Warrior, U., Kofron, J.L., and Burns, D.J., Microarrayed compound screening (μARCS) to identify activators and inhibitors of AMP-activated protein kinase, *J. Biomol. Screen.*, 9, 112, 2004.

94. Burns, D.J., Kofron, J.L., Warrior, U., and Beutel, B.A., Well-less, gel-permeation formats for ultra-HTS, *Drug Discov. Today*, 6, S40–S47, 2001.
95. Taylor, P.B., Stewart, F.P., Dunnington, D.J., Quinn, S.T., Schulz, C.K., Vaidya, K.S., Kurali, E., Lane, T.R., Xiong, W.C., Sherrill, T.P., Snider, J.S., Terpstra, N.D., and Hertzberg, R.P., Automated assay optimization with integrated statistics and smart robotics, *J. Biomol. Screen.*, 5, 213, 2000.
96. Cunningham, B.T., Li, P., Schulz, S., Lin, B., Baird, C., Gerstenmaier, J., Genick, C., Wang, F., Fine, E., and Laing, L., Label-free assays on the BIND system, *J. Biomol. Screen.*, 9, 481, 2004.
97. Kong, Y., Aklian, M., Pandey, N., Lynch, T., and Wildey, M.J., Label-free detection in high-throughput format for cytokine binding interaction and functional kinase assays, presented at *IBC Assays and Cellular Targets Meeting*, San Diego, October 18–21, 2004.
98. Landers, P., Drug industry's big push into technology falls short, *The Wall Street Journal*, February 24, 2004, p. A1.

6

Pharmacological and Pharmaceutical Profiling: New Trends

Joanne Bowes, Michael G. Rolf, Jean-Pierre Valentin, Valérie Hamon, Mark Crawford, and Thierry Jean

CONTENTS

Part 1 is written by Joanne Bowes, Michael G. Rolf, and Jean-Pierre Valentin from AstraZeneca R&D Alderley Park, U.K. and Part 2 by Valérie Hamon, Mark Crawford, and Thierry Jean from Cerep, U.S.A.

Part 1: Pharmacological Profiling: The Early Assessment of Pharmacological Properties

6.1 What Is Pharmacological Profiling?

In this chapter, pharmacological profiling is a term used to describe the investigation of the pharmacological effects of a drug at molecular targets distinct from the intended therapeutic molecular target. It may also be described as selectivity screening or secondary pharmacology. To mediate a biological effect, a drug binds specifically to the primary therapeutic target or to other molecular targets (Figure 6.1). These effects may be mediated by the parent drug or by metabolites of the drug. These molecular targets may be closely related to the therapeutic target (e.g., receptor subtypes) or very distantly related — for example, the therapeutic target may be an enzyme, and profiling investigates activity at other enzymes as well as receptors, ion channels, and transporters.

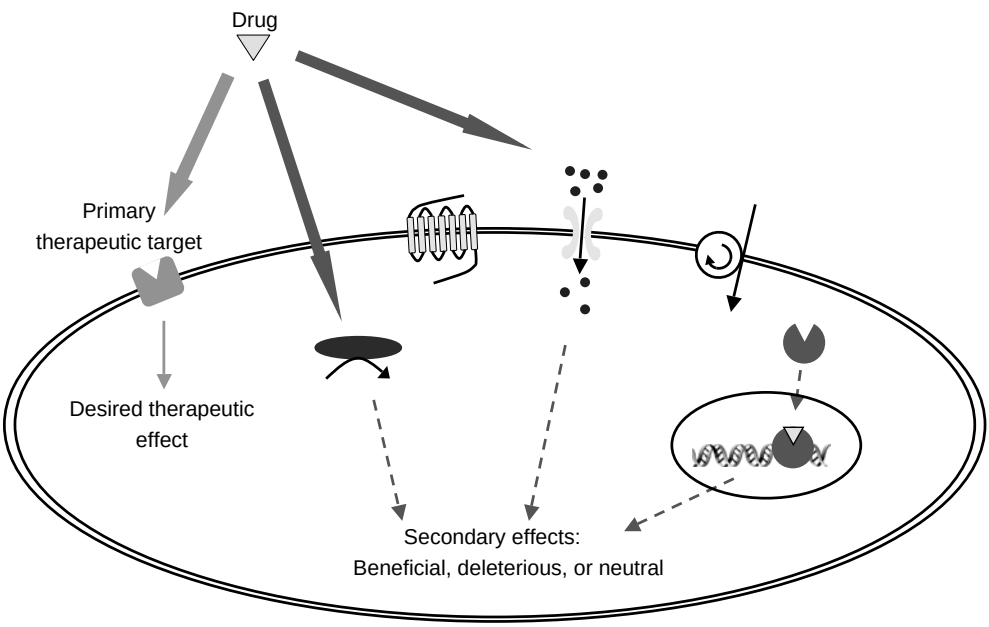


FIGURE 6.1
Mechanisms of drug action. To mediate a response the drug can bind to the desired therapeutic target or to other molecular targets such as G-protein-coupled receptors (GPCRs), ion channels, or transporters on the cell membrane, or to intracellular targets such as enzymes and nuclear hormone receptors.

6.2 Why Do Pharmacological Profiling?

6.2.1 The Regulatory Drive for Safe Medicines

Adverse drug reactions in man account for 10% of the failure of drugs in clinical development.[1] In addition, adverse drug reactions account for ~5% of all hospital admissions and around 0.15% of deaths following admission.[2,3] These adverse drug reactions may be because of the activity of the drug at the primary molecular target, activity at other molecular targets, or may be mediated by nonspecific mechanisms. During the development of a drug, extensive preclinical safety studies are performed to assess the effects of the drug, and the selection and design of these studies is influenced strongly by worldwide regulatory requirements. The International Conference on Harmonisation (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use, is a body of experts that publishes guidelines and recommendations pertaining to the clinical use of new chemical entities (NCEs). Major regulatory authorities worldwide adopt and implement these guidelines. The ICH S7A[4] document outlines recommendations for safety pharmacology studies for human pharmaceuticals and defines three types of pharmacology studies that are applied during the drug discovery and development process. They are defined as follows:

- *Primary pharmacodynamic studies* investigate the mode of action or effects of a substance in relation to its desired therapeutic target.
- *Secondary pharmacodynamic studies* investigate the mode of action or effects of a substance not related to its desired therapeutic target.
- *Safety pharmacology studies* investigate the undesirable pharmacodynamic effects of a substance on physiological functions in relation to exposure in the therapeutic range and above.

Safety pharmacology studies are a regulatory requirement. Pharmacological profiling studies (referred to as "secondary pharmacodynamic studies" by ICH S7A) are recommended, but not legally required. In contrast to most safety pharmacology studies, pharmacological profiling does not need to be conducted in compliance with Good Laboratory Practice (GLP).[4,5] The ICH S7A guidelines clearly indicate that pharmacological profiling should be considered and applied in two ways. The studies should be applied to provide information for the selection and design of safety pharmacology studies (proactive approach) and used to interpret the findings of *in vivo* safety pharmacology and toxicology studies or clinical findings (retrospective approach). Ultimately, it is desirable for a drug to be devoid of any adverse drug reactions when first administered to humans and when used in clinical practice. The end result is the development of a safe medicine for the patient.

6.2.2 Benefit to Drug Discovery and Development

A major benefit of pharmacological profiling is the potential influence on decision-making in drug discovery. If key decisions can be made earlier in discovery, this avoids complications during the later, more resource intensive and expensive phases of drug development. One example of decision-making at early stages is where multiple chemical series are being evaluated and the primary pharmacology and physicochemical properties of the compounds make it difficult to distinguish between one chemical series and another. By

generating a pharmacological profile, the information can influence the decision as to which chemical series is the best to take forward into the next phase of the drug discovery process. The structure–activity relationships can be explored within a chemical series and the unwanted activities designed out, or additional beneficial activities that are identified can continue to be incorporated. Understanding the pharmacological profile of a drug is also useful in interpreting observed functional responses in *in vivo* studies, toxicology studies, or most importantly, in clinical studies or in patients. If the *in vitro* pharmacology data highlight activity at a target that may result in a deleterious effect, then specific *in vivo* studies can be designed to explore the functional effects in relation to the expected therapeutic plasma concentration (taking into account plasma protein binding) of the drug. Greater understanding of the profile of the compound lowers the risk of an adverse drug reaction in volunteers in clinical trials and in patients in clinical practice that cannot be explained.

Pharmacological profiling can also reveal potential beneficial effects: drug discovery by serendipity. In this case, it may be that the drug has dual activity on two targets that may play a role in a disease. For example, a drug that is a dual antagonist at histamine H_1 receptors and platelet-activating factor (PAF) receptors (e.g., rupatadine) has added benefit over a selective histamine H_1 antagonist (e.g., loratadine) for the therapy of allergic rhinitis because both PAF and histamine play key roles in mediating inflammatory responses.[6] Similarly, a drug that has dual activity at the dopamine D_2 receptor and β_2-adrenoceptor (e.g., sibenadet) has the potential for greater efficacy in the therapy of chronic obstructive pulmonary disease when compared with a selective β_2-adrenoceptor agonist (e.g., salbutamol).[7,8]

Pharmacological profiling may also result in the discovery of novel drugs at known targets. For example, by profiling a drug that is being developed as a potent blocker of an ion channel, affinity at a G-protein-coupled receptor (GPCR) may be identified. The drug may represent a novel chemical structure with activity at the GPCR, resulting in the generation of a new patent. If the drug had not been profiled the pharmaceutical company would be unaware of this potentially novel drug for the therapy of, perhaps, another disease indication.

Understanding the pharmacological profile of a drug contributes to the development of drugs that have improved safety and tolerability over existing therapies for a disease and this provides the pharmaceutical company with a significant advantage over their competitors and benefit to the patients. For example, tricyclic antidepressants (TCAs) were discovered in the 1950s–1960s and were the therapy of choice for depression. TCAs (e.g., amitryptyline) nonselectively inhibit the reuptake of serotonin, noradrenaline, and dopamine and have a range of side effects that are a consequence of the lack of selectivity for the monoamine uptake transporters and activity at a range of receptors such as histamine, muscarinic, and adrenoceptors.[9] The therapy of choice for depression nowadays are the selective serotonin reuptake inhibitors (SSRIs) that have similar efficacy, but significantly fewer side effects (e.g., sertraline),[10,11] and therefore increased patient compliance. Another example is the utilization of phosphodiesterase (PDE) 5 inhibitors for the therapy of erectile dysfunction.[12] For example, tadalafil has equivalent efficacy but is more selective for PDE 5 over PDE 6 (780-fold selective for PDE 5)[13] when compared with sildenafil that has the potential to inhibit PDE 6 (approximately tenfold selective for PDE 5 over PDE 6) at higher doses. Inhibition of PDE 6 is associated with visual side effects in the form of disturbance in discrimination between the colors green and blue and increased perception of brightness mediated by inhibition of PDE 6 activity in the retinal rod cells of the eye.[14,15] The greater selectivity of tadalafil for PDE 5 results in potentially a safer drug profile for the patient and a competitive edge for the pharmaceutical company that developed it.

6.3 Methods Applied to Pharmacological Profiling

Drugs exert effects by binding to biological protein targets that may be receptors, enzymes, ion channels, or transporters. The pharmacological action of drugs can be investigated *in vitro* or *in vivo* by a variety of methods, all of which are aimed at quantifying the drug action at a particular target so as to compare one drug with another. The parameters measured can be (i) direct binding affinity of a drug to a target, (ii) a functional response in a cell or tissue, and (iii) a physiological response in an animal. A combination of these studies allows an integrated assessment of the biological effects (pharmacology) of a drug. This section will focus on methods applied to pharmacological profiling that are applicable to profiling small molecule drugs as opposed to biological agents.

6.3.1 Radioligand-Binding Assays

Pharmacological analysis in the simplest form for receptors, ion channels, and transporters is determination of the affinity of a drug for a target and this is determined by performing *in vitro* radioligand-binding assays. These assays are designed to measure the ability of a drug to compete with a radiolabeled compound that is selective for a particular target. Concentration–response curves are constructed and activity quantified as IC_{50}: the molar concentration of drug that inhibits the binding of a radioligand by 50% (Figure 6.2).[16] If the binding is competitive this can be converted into a K_i value using the Cheng–Prusoff equation.[17] K_i is the inhibition constant for a drug: the concentration of competing ligand in a competition assay that would occupy 50% of the receptors if no radioligand was present. The IC_{50} value for a drug may vary between experiments depending on radioligand concentration used, whereas K_i is an absolute affinity value.

There are many advantages when applying the technique of radioligand-binding assays to pharmacological profiling; they are technically easy to perform, only small quantities of drug are required, and if validated correctly they are extremely robust, reproducible, and suitable for high-throughput screening (HTS) formats. Importantly, the drug is incubated

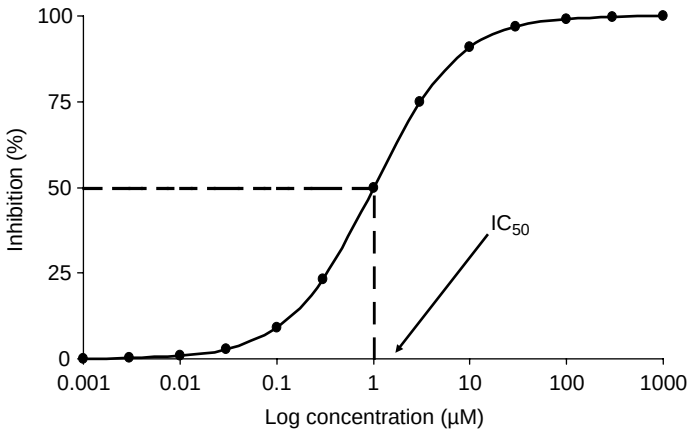


FIGURE 6.2
An example of a concentration–response curve for a drug in a radioligand-binding assay. Increasing concentration of drug (*x*-axis) causes an increase in the percentage inhibition of binding of the radiolabeled compound. The IC_{50} is defined as the concentration of drug that displaces 50% of the specific binding of the radioligand. In this example, the IC_{50} for the drug is 1 μM.

to reach equilibrium so quantitative determinations of affinity are attainable. Where the target has been cloned, the drug can be profiled against the human isoform of the target by screening using membranes from cells that express the recombinant form of the human receptor. Where in-house resources are not available, there are contract research organizations that specialize in the technique of radioligand binding for pharmacological profiling. However, there are some limitations to this technique. The main restriction is that binding assays do not provide any information on the pharmacological mechanism of action of the drug, i.e., they cannot distinguish between agonists, antagonists, partial agonists, inverse agonists, or allosteric modulators. In addition, radioligand-binding assays may be useful, but are not optimal for assessment of drug activity at all targets. For example, they may be inferior to functional electrophysiology assays for ion-channel targets, and the use of recombinantly expressed proteins means that assumptions are made that the binding affinity will be equivalent at the native human target. Overall, the interpretation and prediction of the functional consequences of a drug having affinity for a target in a radioligand-binding assay is a challenge, and the relevance of radioligand-binding assay data may only be revealed when a drug is subsequently tested in functional *in vitro* and *in vivo* assays and most importantly in humans. This will be discussed below. Many of the points discussed apply to enzymes where the most straightforward approach is to perform simple *in vitro* enzyme activity assays. Instead of measuring the ability of the drug to displace a radiolabeled compound, the ability of the drug to compete with the turnover of a substrate into a product is measured. Potency is determined in the same way as for radioligand-binding assays, i.e., expressed as IC_{50} or K_i. The substrate may be radiolabeled or in the form of a colorless substrate that is cleaved to form a colored product, in a reaction that can be easily quantified.

6.3.2 Functional *In Vitro* Assays

Pharmacological analysis of a drug in functional assays involves determining whether it is an agonist or antagonist at a particular receptor, and the measurement of the potency of the drug at the receptor. An agonist is a drug that binds to a receptor and activates it to produce a biological response, and its effects depend on affinity (tendency to bind to a receptor) and efficacy (ability once bound, to initiate changes which result in the biological effect).[16] An antagonist is a drug that binds to a receptor but does not result in a biological response (i.e., has no efficacy), but attenuates the response of an agonist. Antagonists can be competitive antagonists that shift the concentration–response curve of an agonist to the right without affecting the maximum response of the agonist, or noncompetitive antagonists that shift the concentration–response curve of an agonist to the right and decrease the maximum response of the agonist. Potency is a measure of the concentration of a drug at which it is effective. For agonists this is expressed as EC_{50}: the molar concentration of an agonist that produces 50% of the maximum possible effect of that agonist.[16] Antagonist potency is often expressed as pA_2 or pK_B. A pA_2 is the negative logarithm to base 10 of the molar concentration of antagonist that would produce a twofold rightward shift in the concentration–response curve for an agonist.[18] A pK_B is a measure of the potency of a competitive antagonist: the negative logarithm to base 10 of the molar concentration at which equilibrium would occupy 50% of the receptors in the absence of agonist. In some cases, it is acceptable to determine an IC_{50} in a functional assay providing an appropriate agonist concentration to be used (i.e., a submaximal concentration).

Functional *in vitro* assays can be whole-cell assays or animal tissue assays. For *in vitro* cell-based assays, recombinant molecular targets can be expressed in cell lines and linked to second-messenger systems to measure functional responses. Examples of functional

readouts in cell-based assay include second messengers such as measurement of changes in intracellular calcium using fluorescent dye-based assays.[19,20] Cell-based assays have many advantages in that they are amenable to medium-throughput screening, a single functional response is measured, and quantitative data are generated. Importantly, these assays provide key information as to whether the drug is an agonist or antagonist at the target. Cell-based assays also reduce the use of animals for these studies and have the added advantage that the drug is profiled at human targets, often in human cell systems. The main limitations are that cell culture on a relatively large scale is resource-intensive, and in some assays true equilibrium may not be achieved.[19] An alternative method used for *in vitro* functional assays is the use of animal tissue — the more traditional "organ-bath pharmacology" to measure physiological responses in an intact piece of tissue. Some examples include investigation of the muscarinic-receptor-mediated contraction of guinea-pig ileum, of β_1-adrenoceptors in mediating the force and rate of contraction of rat-isolated atria and α-adrenoceptors in mediating vasoconstriction of rabbit aorta.[21] The advantage of these assays compared with cell-base assays is that the drug can be incubated over time to ensure equilibrium, thereby enabling potency to be determined quantitatively. The main limitations of this type of functional assay are the use of tissue from animals, and that the drug is profiled at nonhuman targets. This may result in incorrect interpretation and prediction of functional effects of the drug in clinical studies if the potency and efficacy at the human and animal target are different. However, in some cases, functional assays can be performed using human tissue, although this may be limited by availability of human tissue for these experiments.[22]

6.3.3 *In Vivo* Pharmacology Studies

The pharmacology of a drug may also be investigated by administering the drug to animals and measuring biological responses. The responses can be pharmacologically quantified by expressing the response as an ED_{50}: the dose that gives 50% of the maximum response to that drug. This type of study may be applicable, for example, in the investigation of effects of the drug on a major physiological system, such as the cardiovascular system. In this case, the drug may be administered to either conscious or anesthetized animals and its effects on heart rate, blood pressure, and regional blood flows can be explored over a range of doses. To interpret the effects, it is useful in these studies to measure the concentration of drug in the plasma when the effect is observed. The advantage of this approach is that the effect of the drug is determined in an intact integrated physiological system. The major limitation in this methodology is that the receptor mediating the response is not known, as many different targets may mediate the same response, which adds complexity to the interpretation. Also, subtle effects may not be detected in *in vivo* studies that are primarily designed to measure certain parameters. Another limitation of *in vivo* studies is the use of animals, which ideally should be kept to a minimum.

6.4 Strategy for Pharmacological Profiling

It is likely that there is considerable diversity in pharmacological profiling strategies adopted by different pharmaceutical companies. The ultimate objective is to establish the selectivity of the drug and identify or predict pharmacological effects that may arise as a consequence of drug action at these targets. For a strategy to have impact on drug discovery and development, it has to be applied in a consistent way within a company to maximize the output and analysis of the data.

One strategy for pharmacological profiling is to front-load *in vitro* screening at key points in the drug discovery process. Pharmacological profiling is applied in a stepwise process where the drug is assayed for activity at an increasing number of molecular targets as it progresses through the drug discovery process (Figure 6.3). The screens applied in the early stages of discovery consist of a small panel of key targets (e.g., 20) that mediate functional effects of the vital physiological systems: cardiovascular, central nervous, and respiratory systems. During the optimization phases, the profile is expanded to a greater number of assays (e.g., 50 to 150) adding targets that mediate effects on other organ systems such as the gastrointestinal system. As the drug evolves into a candidate for the development phase, a greater number of targets (e.g., up to 300) covering all organ systems are screened. The result is the generation of a comprehensive pharmacological profile of the drug before it is first administered to humans.

The assay methodologies applied are firstly to perform *in vitro* radioligand-binding assays and enzyme activity assays, using human targets wherever possible. The drug is initially screened at an appropriate single concentration and any significant hits (significant is defined as >50% inhibition) are retested to determine IC_{50} or K_i values to provide a quantitative analysis of the activity (Figure 6.4). The test concentration is selected on the basis of the affinity of the drug at the therapeutic target, and it is recommended that the test concentration should be a multiple of the primary therapeutic target IC_{50}, to use the data to assess the therapeutic margin. For example, the test concentration may be tenfold higher for drug in the early phases of discovery that are not yet optimized, but for drugs in the later phases it may be appropriate to test at 100-fold higher than the therapeutic target potency. Once the activity of the drug at the other molecular target has

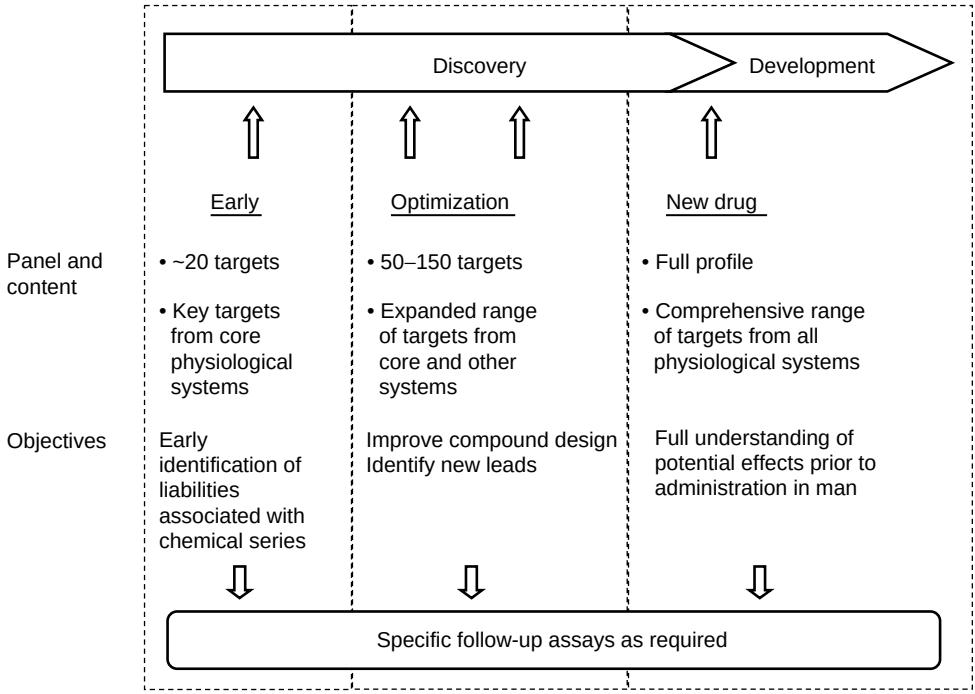


FIGURE 6.3
Pharmacological profiling in the drug discovery and development process. The number of molecular targets tested increases at each stage of the process and the information generated at each stage assists different types of decisions. The strategy is on the basis of a stepwise approach, focussed on front loading of *in vitro* assays to generate a comprehensive pharmacological profile prior to the first administration of the drug into humans.

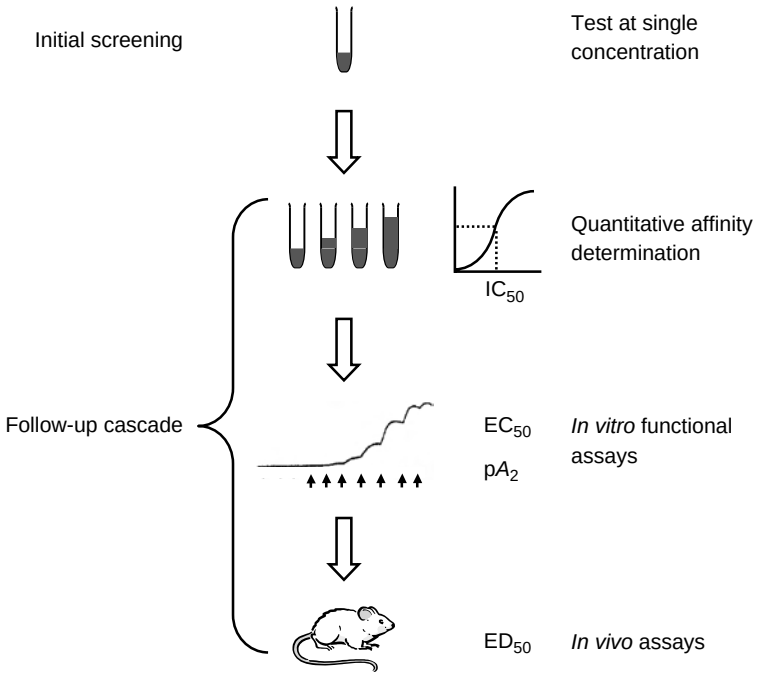


FIGURE 6.4
The stepwise process applied to pharmacological profiling. Initially *in vitro* radioligand-binding assays or enzyme activity assays are used to profile the drug. If the drug has significant (>50% inhibition) activity, then *in vitro* functional assays are performed to understand the mode of action of the drug at the target and, if required, *in vivo* studies are subsequently performed.

been quantified in a radioligand-binding assay, the mode of action is explored in *in vitro* functional assays to determine whether the drug is an agonist or antagonist at that target. Functional effects in relation to dose and plasma exposure of the drug can then be explored in specifically designed *in vivo* studies (Figure 6.4).

An example of the application of this stepwise approach (*in vitro* screening followed by *in vivo* studies) is the investigation of the drug RS-15385-197.[23] The first step involved profiling the drug in radioligand-binding assays (16 molecular targets) and the drug was found to be selective for α_2-adrenoceptors with a pK_i of 9.5. The next step was to investigate the effects of the drug in *in vitro* functional tissue assays. RS-15385-197 antagonized the effects of the α_2-adrenoceptor agonists UK-14,304 and BHT-920, in the transmurally stimulated guinea-pig ileum (pA_2=9.7) and the dog saphenous vein (pA_2=10.0), respectively. The following step involved *in vivo* studies where RS-15385-197 antagonized the mydriasis response induced by the α_2-adrenoceptor agonists UK-14,304 or clonidine and the pressor response to UK-14,304 in pithed rats. So, in this example, initial activity in the radioligand-binding assay translated into significant *in vivo* effects in rodents.

6.5 Impact of Pharmacological Profiling

The impact of the *in vitro* screens applied to each stage in the drug discovery process will be different. The optimal set of assays at the early stages should assess drug activity against a small number of targets that play a key role in the function of core physiological systems.

A well-designed set of targets facilitates decision-making, and the interpretation relating to these targets is straightforward. For example, the panel may include key receptors that control the autonomic nervous system such as adrenoceptors and muscarinic receptors, and other key targets known to cause adverse drug reactions in humans (some examples are given in Table 6.1). As a drug progresses further along the drug discovery process, the data may be used to identify unwanted activities and this would allow the medicinal chemists to explore the structure–activity relationships at the molecular target and design out these undesirable activities from the chemical series. As the drug evolves into a potential new drug for development, and when preclinical safety pharmacology and toxicology studies are being considered, pharmacological profiling contributes to the study design by identifying potential safety issues that will need to be monitored. Dosing levels may also be set, taking into account the pharmacological profile, so that sufficient exposure is achieved in the studies to investigate the potential effects mediated by other molecular targets.

6.6 Predicting Functional Responses *In Vivo* from *In Vitro* Data

Having generated the *in vitro* data to explore the pharmacological profile of the drug, the next step is to interpret the data and make predictions of the potential functional effects as a result of a drug having this profile of activity. There are certain factors that should be considered. First of all, the affinity of the drug at other molecular targets should be analyzed relative to the affinity at the desired therapeutic target and to the human predicted (or measured) maximum free plasma concentration of the drug to assess therapeutic margins. The data can be presented in the form of a "receptogram" (Figure 6.5). The general rule of thumb is, if the affinity at the other molecular targets is 100-fold lower than the affinity at the desired therapeutic target then it is unlikely that there will be any observed functional effects in humans, as it is unlikely that the drug will reach levels high enough to have significant activity at the secondary target. However, this "margin" may differ depending on

TABLE 6.1

Examples of Functional Effects Due to Activity at Pharmacological Targets

Target	Mechanism	Administration (Site of Action)	Potential Functional Effect	Example
PDE4	Inhibitor	Acute (central nervous system)	Emesis	Rolipram[24,25]
Muscarinic M_3 receptor	Antagonist	Acute (gastrointestinal system)	Reduced GI motility	Darifenacin[26]
Histamine H_1 receptor	Antagonist	Acute (central nervous system)	Sedation	Diphenhydramine[27]
β_1-adrenoceptor	Agonist	Acute (cardiovascular system)	Tachycardia	Isoprenaline[28]
Cyclooxygenase-1	Inhibitor	Chronic (gastrointestinal system)	Ulceration/ bleeding in gastrointestinal tract	Naproxen[29]
hERG potassium channel	Blocker	Acute (cardiovascular system)	QT prolongation, cardiac arrhythmia	Cisapride[30]
hERG potassium channel	Blocker	Chronic (reproductive system)	Teratogenicity	Cisapride[31]

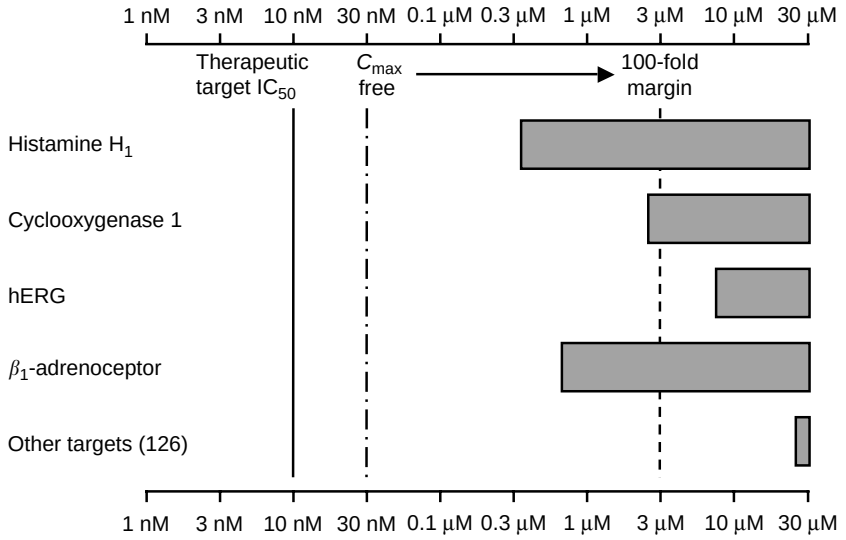


FIGURE 6.5
An example of a "receptogram." This diagram can be used to assess the therapeutic margins by plotting the IC_{50} values for the targets and comparing these to the therapeutic target affinity and predicted free plasma levels for therapeutic efficacy. In this example, the drug shows selectivity of less than 100-fold at three molecular targets when tested in a panel of 130 screens. The potential functional effects of the drug because of the actions at these targets are described in Table 6.1

the target under analysis. For the human ether-a-go-go related gene (hERG) potassium channel, it is clear that to be confident that proarrhythmia will not be observed in the clinic, the margin between the potency of the drug at the therapeutic target and potency at hERG should be at least 30-fold.[32] However, being realistic, data are not currently available in the public domain for many of the targets explored in pharmacological profiling, and so the generalized "100-fold" margin may need to be reviewed in the future as more data become available. The next point to consider is the location of expression of the target. For example, targets primarily expressed in the brain will not be activated *in vivo* if the drug does not cross the blood–brain barrier, and so side effects potentially mediated by activity of the drug at these receptors can be ruled out under normal physiological conditions. Analysis of the physicochemical properties of the drug allows predictions to be made about the ability of the drug to penetrate the brain. Conversely, if a drug is administered by the oral route, the local concentration of the drug in the gastrointestinal tract may be high relative to the plasma efficacious concentration, and functional effects as a consequence of activity at a target expressed in the gastrointestinal tract may be revealed. Alternatively, the drug may be developed for administration by inhalation, where the focus for interpretation will be on targets expressed in the respiratory system, and it is unlikely that there will be significant systemic exposure of the drug. Another point to consider is the different interpretation that may be applied if a drug is administered chronically (Table 6.1). For example, drugs with activity at the β_2-adrenoceptor (e.g., isoprenaline) or the K_{ATP} channel (e.g., minoxidil) cause peripheral vasodilation and this is associated with reflex tachycardia.[33] However, there is clear evidence that chronic administration of these drugs results in necrosis of the left ventricle and papillary muscle of the heart in dogs.[33,34] Interestingly, this effect does not seem to translate into humans.[35] The reason for this is that in humans, heart rate is closely controlled, whereas in dogs these compounds elevate heart rate to 180 to 200 beats/min and it is this high heart rate that results in necrosis. Identification of vasodilator action from

pharmacological profiling can be used to predict and explain this observed pathology but provide confidence that this is unlikely to translate into humans.

Another key question is whether the affinity of a drug determined in a radioligand-binding assay correlates with affinity determined in a functional assay. For some receptors, the affinity estimated from binding and isolated tissue functional assays are equivalent, for example muscarinic receptors[36] and α-adrenoceptors,[37] but for other targets there are differences. For example, binding K_i values for some β_2-adrenoceptor agonists do not correlate with potencies in isolated tissues[38] — this is more likely for agonists rather than antagonists because agonism depends on affinity and intrinsic activity and only affinity can be measured in binding assays. Other factors affect the correlation of *in vitro* binding data with *in vitro* functional tissue assays for example, the assay conditions (e.g., pH, temperature, protein concentration, equilibration time), the choice of preparation (e.g., membrane vs. whole-cell binding assay), the species selected (human vs. rodent), the drug metabolism capacity of the tissue, and the partitioning of the drug into the tissue.

The next step is to investigate whether the *in vitro* functional effects are revealed *in vivo*. In some cases, the target mediates a physiological response and will clearly be revealed *in vivo*, for example, inhibition of PDE4 activity is associated with emesis (rolipram[24,25]), antagonism of the histamine H_1 receptor in the CNS causes sedation (diphenhydramine[27]), and inhibition of PDE6 activity causes visual disturbances (sildenafil[12]). However, it is also possible that compensatory mechanisms may override the drug's action and a net functional effect may be undetectable. For example, if a drug has activity at the hERG potassium channel and also at the L-type calcium channel, although antagonism of the hERG potassium channel may result in the prolongation of the QT interval of the electrocardiogram, blockade of the L-type calcium channel may result in the shortening of the QT interval. In this case, it may be anticipated that overall there would be no effect on the QT interval *in vivo*. This is the case for verapamil, which blocks both L-type calcium channels and hERG potassium channels, and does not have any effect on the QT interval *in vivo*.[39]

6.7 Maximizing the Strategy: Generation of Databases

To maximize the impact of a pharmacological profiling strategy, the application of a database can serve as a good tool to aid in the analysis of data and to capture target-related interpretation. The identification of "false-positive" data (apparent activity in a binding assay that is not confirmed *in vivo*) or "false-negative" data (where an *in vitro* assay fails to pick up activity observed *in vivo*) is a key aspect of pharmacological profiling. Application of the described strategy in a consistent way and collating all of the *in vitro* and *in vivo* data into a database acts as a good mechanism to improve the detection of false positives or false negatives. Data from assays exhibiting unusual hit rates can be compared to results from functional *in vitro* or *in vivo* assays and false positive or negative results identified. This serves as a mechanism for quality control checks and assay protocols can be reexamined and optimized, or replaced, if necessary. The database should also include other data relating to each molecular target, such as binding affinities for standard reference compounds, pathways linked to the activation or inhibition of targets, clinical trial outcomes, and reported adverse drug reactions. These data can be extracted from the published literature and it is hoped that in the future more data will become available in the public domain. Relating these data to the results of pharmacological profiling experiments provides a powerful tool to aid in the interpretation of results. Where many

compounds have been profiled *in vitro* and also *in vivo* in animals or humans, linking these data together in a database also helps to assess the predictivity of *in vitro* assays for clinical outcomes, the ultimate purpose of pharmacological profiling.

6.8 *In Silico* Pharmacological Profiling

As well as facilitating the interpretation of profiling data, construction of a database capturing profiling data and supplementary information opens up new possibilities for further extending the pharmacological profiling strategy. In the future, profiling may involve computer-based models in place of laboratory assays. Computational models have been built for some molecules related to the action of the drug, such as the hERG potassium channel, cytochrome P450 enzymes, and other promiscuous proteins.[40] Compared to the large number of potential molecular targets, the number of published models is still very small. This may be partially explained by the many challenges to be overcome when assembling a dataset that enables predictive work *in silico*.[41] However, if a consistent strategy is implemented then many of the historical issues surrounding *in silico* modeling, such as inconsistent datasets, can be avoided. Several companies have expended considerable resources creating proprietary databases focusing on different areas of pharmacological profiling (see Table 6.2). The systems have many features in common — datasets including large numbers of marketed drugs profiled against large panels of *in vitro* assays, and supplementary compound and target-related information. Access to these tools is generally provided on a license basis or *via* collaborative efforts between a pharmaceutical company and the database provider. These arrangements reflect the developmental status of such systems and the high startup costs associated with compiling the data. *In silico* pharmacological profiling can be applied in many ways, to predict activity of a drug at pharmacological targets without the need to do actual experiments. A simple first step involves comparing the structure of a new drug to structures of drugs that have already been profiled. Examination of the profiles of similar compounds may reveal common pharmacological activities and this may aid in the identification of activities that may be detected *in vivo*. This can be done quickly and easily. A second, more advanced strategy involves building detailed pharmacophore models for the targets contained within the database, so that predictions can be made for each individual target. In theory, this could then allow a full pharmacological profile to be generated without actually needing to run laboratory

TABLE 6.2

Pharmacoinformatics Providers

Product	Company	Comment
Bioprint™	Cerep [http://www.cerep.fr/]	Contains extensive data on G-protein-coupled receptors (GPCRs), targeted toward drug discovery
DrugMatrix™	Iconix Pharmaceuticals, Inc. [http://www.iconixpharma.com/]	Pharmacology data integrated with extensive gene expression data
Receptor Selectivity Mapping Database (RSMDB®)	Novascreen Biosciences Corporation [http://www.novascreen.com/]	Focuses on optimizing lead series design

assays. This would have significant benefits in terms of the number of compounds that could be screened in a given time, using minimal resources. Another consideration is the impact on the 3Rs rule of animal welfare (reduction, refinement, and replacement), which is particularly important in some countries. However, current computational pharmacophore models are generally not accurate enough to act as the sole source of pharmacological data. While a model may be sufficient to rank several compounds within a chemical series, absolute affinity or potency estimates are often not reliable and predictions based on a single compound can be misleading. Therefore, *in silico* models are currently best used as a high-throughput filter allowing large numbers of chemical structures to be screened at low cost. The most promising candidates would then be further profiled using the experimental pharmacology assays. Data on compounds with *in silico* predictions that have been validated by laboratory assays can then be used to further improve the *in silico* models. In the future, *in silico* models may replace *in vitro assays*, at least for early-stage screening of drugs prior to regulatory pharmacology and toxicology studies.

6.9 Concluding Remarks

Adverse drug reactions are a major cause of failure for drugs entering clinical development. Marketed drugs with undesirable side effects are a major factor in reducing patient compliance. Every drug that enters clinical development represents an investment of millions of dollars that will only be recouped if the drug is approved for use and is frequently prescribed. Therefore, it is extremely important to ensure that new drugs are safe and tolerable as well as efficacious. By introducing pharmacological profiling early in the drug discovery process and increasing the scope of profiling as drugs progress toward the clinic, a pharmaceutical company can minimize the risks associated with unwanted side effects. Early *in vitro* profiling, utilizing an incremental approach developing in parallel with a drug's progression through the stages of drug discovery represents the optimal approach to profiling, supplying high-quality data at key decision points. These activities may also result in the discovery of novel drugs at known targets; discovery by serendipity. Consistent application of this strategy within a pharmaceutical company's portfolio is the key to success.

Part 2: Strategic and Technical Aspects of Drug Profiling

6.10 Introduction

Pharmacology developed as a science to characterize biological properties of natural and synthetic compounds. Initially, biological properties were revealed through testing in animal models or at the best in animal tissues. Progress in physiology, and cellular and molecular biology allowed the development of *in vitro* models that were somewhat predictive of *in vivo* properties. Pharmacology then evolved toward the testing of an increasing number of compounds. In the 1990s, genomic and molecular biology provided scientists with thousands of new proteins to be considered as potential targets for drug discovery. Combinatorial chemistry and HTS became the key technologies for an innovative drug discovery approach

based on large numbers. Through a simplistic calculation, multiplying the number of targets (a few thousand) by the number of compounds to screen (in the range of millions in large pharmaceutical firms), motivated the development of ultra-HTS technologies. Indeed, an increasing number of compounds active on new targets were discovered and chemically optimized. However, these efforts and successes in early drug discovery programs did not result in an increased number of NCE. The number of NCE in 2002 and 2003 were lower than any other 2-year period in the preceding 10 years.[42,43] If more compounds are synthesized, screened, and optimized, whereas in parallel fewer new drugs reach the market, it means that more drug candidates are failing in later discovery and development steps.

Poor drug efficacy, toxicology, safety, and physicochemical characteristics are responsible for most drug failures. Profiling developed recently as a method of predicting, through the use of *in vitro* assays, the behavior of lead candidates in man and aid in reducing failures in clinical trials. Drug profiling provides information of high value to support the design and the selection of drug candidates with higher probability of success. To be efficient, drug profiling must be done in a battery of well-selected *in vitro* assays, i.e., assays that are robust and predictive.

This part of the chapter will focus on both strategic and technical aspects of drug profiling. We will review the criteria to be taken into consideration for the selection of assays to include in a profile and discuss the new trends in profiling.

6.11 Pharmacological Profiling as the Tool to Avoid Drug Failure

6.11.1 Economic Environment: The Cost of Drug Failure

Discovery and development of new drugs require money, time, and luck. Using current methods, the average drug development time from target identification to New Drug Application (NDA) filing is 12 years.[44,45] Research and Development (R&D) expenses have kept increasing since the 1970s, reaching up to 13% of revenue.[46,47] Despite this increasing effort, the number of NCE reaching the market has been declining since 1990 with only 17 NCE launched in 2002.[48] Bringing a new chemical entity to market has an average cost of more than US $800 million.[49] Interestingly, 70% of this amount is spent on drugs that fail during discovery or development phases. Considering all therapeutic areas, the average success rate of a drug candidate entering clinical Phase I is 11%.[48]

The expectation that HTS and combinatorial chemistry could by themselves improve the drug discovery process is now clearly erroneous,[50] and effort is currently focussed on developing methods to decrease clinical attrition through early selection of drug candidates with the best chances of success.

6.11.2 Understanding the Causes of Drug Failures

The first step to increase success in clinical development is certainly to understand the causes of drug failure. The major reasons for failures of 198 NCE in clinical development have been reported to be bioavailability (39% of failures), lack of efficacy (30%), toxicity (11%), and adverse effects (10%).[51,52]

More recent studies showed that the relative importance of major causes of drug failures in clinical trials has evolved significantly from 1991 to 2000.[42] In the early 1990s, major causes of clinical attrition were poor pharmacokinetics (PK) and bioavailability which together

were responsible for over 40% of failures.[52] In 2000, these causes contributed less than 10% to attrition in drug development,[42] probably as a result of the efforts made by pharmaceutical firms during this period to treat these issues earlier in the drug discovery phases. The contributions of both safety and toxicity issues to attrition have increased during the decade and represented together more than 30% of the causes of drug failures in 2000.[42]

Even after a drug has been approved and reaches the market, adverse reactions may occur and emerge that have a tremendous impact on patients and the industry. Each year in the United States, 2.2 millions people (7% of patients) suffer from drug side effects and 106,000 of them die. Death from side effects of drugs is the fourth cause of avoidable deaths.[53] As a consequence, the regulatory agencies are demanding in-depth information and evidence related to drug safety.[45] Adverse effects emerging only after marketing are of great concern to pharmaceutical companies as illustrated recently by the market withdrawal of cerivastatin (Baycol, Bayer, Pittsburgh, PA, U.S.A.), valdecoxib (Bextra, Pfizer, NY, U.S.A.), and rofecoxib (Vioxx, Merck, Whitehouse Station, NJ, U.S.A.).

It is a common goal to improve the discovery process through the constant effort, to identify as early as possible, drug candidates that will have minimal chances to fail before or after they are approved for marketing. This can be achieved by the early identification of candidates that will have the best safety profiles, which means drugs that are active on their therapeutic target and have minimal adverse effects.

6.11.3 Predicting Side Effects to Prevent Drug Failures

6.11.3.1 Role of Pharmacological Profiling

Adverse effects may result from the action of the drug on its principal therapeutic target. In that case, there is nothing the chemist can do to optimize the safety of the drug, since any modification of the molecule to increase its affinity and selectivity for the target will increase both its therapeutic and its undesired effects.

It is, however, more frequent that side effects are because of an action of the molecule on targets distinct from that responsible for its therapeutic action. One then wants to optimize toward compounds that are highly selective and specific for their therapeutic target(s). This introduces the concept of pharmacological profiling, i.e., the *in vitro* testing on targets (proteins) involved in potential side effects. Whether pharmacological profiling should be taken into consideration early or late in the optimization process is a question that will be addressed below. Irrespective of the answer to this question, the composition of the profile, i.e., the list of *in vitro* targets against which the drugs will be tested is of great importance.

6.11.3.2 Design of Pharmacological Profiles

A pharmacological profile must contain targets closely associated with the primary target of interest to assess selectivity. As an example, if one wants to discover a ligand specific for the 5-HT2c receptor, it may seem *a priori* logical to include other serotonin receptors subtypes as well as the serotonin transporter in the selectivity profile. In other words, the profile should include targets known to bind serotonin. Experience shows that 5-HT2c ligands cross-react with other nonserotonin-related targets, with sometimes a much higher affinity than with serotonin receptors. How can we identify these targets?

One simple approach would be to select targets that are phylogenetically related to the primary target, with the hypothesis that the amino-acid sequence of a protein is the main factor that controls interaction with ligand. Therefore, two proteins with high degree of sequence homology should exhibit similar affinity for a given ligand. Such a phylogenetic analysis applied to our example of 5-HT2c-related targets is illustrated in Figure 6.6.[54] It

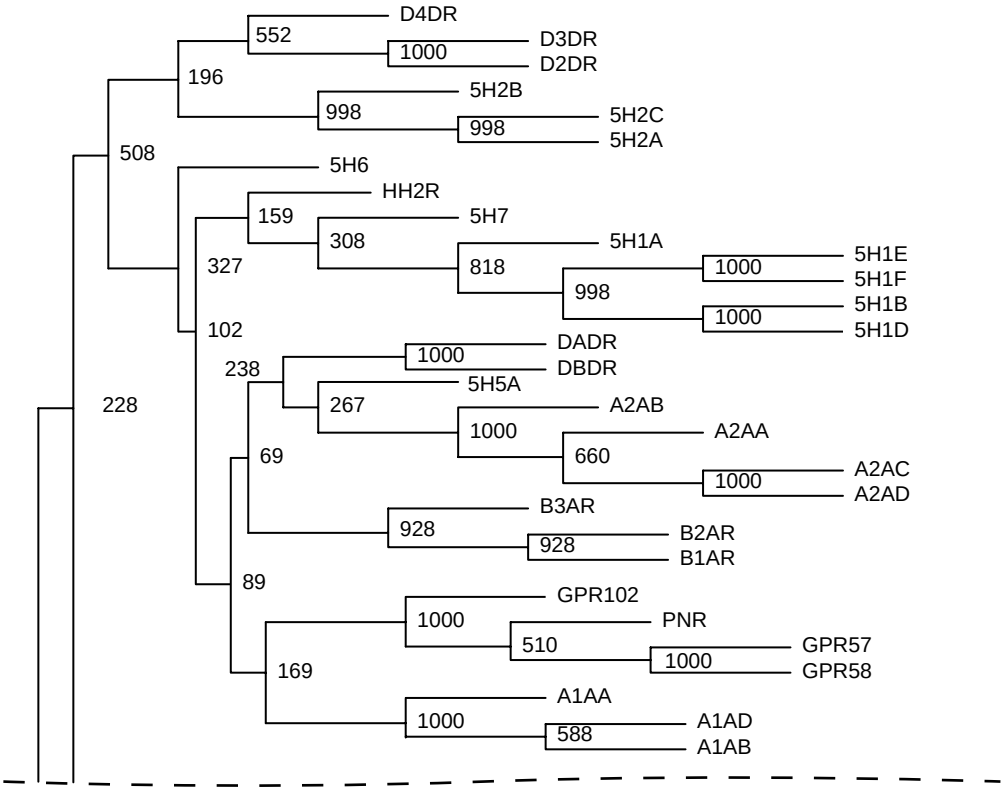


FIGURE 6.6
GPCR phylogenetic analysis: partial neighbor-joining tree of the rhodopsin receptorlike family (A) from Joost and Methner (*Genome Biol*., 3: research 0063.1-0063.16, 2002. http://genomebiology.com/ 2002/3/11/ research/0063). Detailed receptor names are provided in the additional data file 1 of the reference.

shows that receptors for nonserotonin biogenic amines "cocluster" with the 5-HT2c (5H2C) receptor. Thus, dopamine receptor subtypes such as D2 (D2DR), D3 (D3DR), and D4 (D4DR) receptors are closer phylogenetically to 5-HT2c receptor than the 5-HT1 receptor subtypes (5H1X).

These receptors are certainly to be included in the specificity profile for a 5-HT2c ligand. Is this sufficient? This phylogenetic analysis excludes non-GPCR serotonin receptors such as the 5-HT3 receptor and even more importantly the 5-HT transporter. If proteins so dissimilar in their DNA sequence exhibit high affinity for the same molecule (namely serotonin), it suggests that phylogenetic relationships cannot be taken as a unique criteria for the selection of targets to include in a profile aimed to define the specificity of a lead candidate. However, phylogenetic clustering of targets as a first basis to build profiles is useful. A profile based on this approach and the methodology to use it is described in Table 6.3.

Beside a molecular rational to design pharmacological profiles, there is more often an experience-based approach. The pharmacological literature has numerous references associating specific receptor or enzyme interactions with adverse effects. A few selected targets and corresponding physiological effects are presented in Table 6.4. In addition, there is a small but growing literature characterizing the cross-activity of drugs on several targets. Early identification of these off-target effects is critical as illustrated by the elucidation of the role of human ether-a-go-go related gene (hERG) in the phenomenon known as QT prolongation or *Torsade de Pointes*.[55] It should be noted that activity across

TABLE 6.3

Example of Broad Profiling against GPRC Targets Built from Phylogenetic Analysis

Phase 1 Pharmacology Family	Phylogenetic Family Primary Targets	Receptor	Phase 2 Phylogenetic Family Secondary Targets	Receptor
Adenosine			A18	A_{2A} (h)
	A18	A_{2A} (h)	A18	A_1 (h)
			A18	A_{2B} (h)
	A18	A_3 (h)	► A18	A_3 (h)
Adrenergic	A17	α_{1A} (h)	A17	α_{1A} (h)
			A17	α_{1B} (h)
			A17	α_{2A} (h)
	A17	α_{2A} (h)	A17	α_{2B} (h)
			A17	α_{2C} (h)
	A17	β_1 (h)	A17	β_1 (h)
			A17	β_3 (h)
	A17	β_2 (h)	A17	β_2 (h)
			A17	β_3 (h)
Cannabinoid	A13	CB_1 (h)	A13	CB_1 (h)
			A13	CB_2 (h)
Dopamine	A17	D1 (h)	A17	D1 (h)
			A17	D5 (h)
	A17	D3 (h)	A17	D3 (h)
			A17	D2S (h)
			A17	D4.4 (h)
Histamine	A18	H_1 (h)	► A18	H_1 (h)
	A17	H_2 (h)	► A17	H_2 (h)
	–	H_3 (h)	► –	H_3 (h)
Melanocortin			A13	MC_4 (h)
	A13	MC_4 (h)	A13	MC_1
			A13	MC_3 (h)
			A13	MC_5 (h)
Melatonin	A09	ML_2 (MT_3)	A09	ML_2 (MT_3)
			–	ML_1
Muscarinic			A18	M_1 (h)
	A18	M_1 (h)	A18	M_2 (h)
			A18	M_3 (h)
			A18	M_5 (h)
			A18	M_4 (h)
	A18	M4 (h)	A18	M_2 (h)
			A18	M_3 (h)
			A18	M_5 (h)
Neurokinin			A09	NK_1 (h)
	A09	NK_1 (h)	A09	NK_2 (h)
			A09	NK_3 (h)
Opiate	A04	κ (KOP)	A04	κ (KOP)
			A04	δ_2 (h) (DOP)
	A04	μ (h) (MOP)	A04	μ (h) (MOP)
			A04	δ_2 (h) (DOP)
Serotonin	A19	5-HT_{1A} (h)	A19	5-HT_{1A} (h)
			A19	5-HT_{1D}

(*Continued*)

TABLE 6.3 (*Continued*)

Phase 1 Pharmacology Family	Phylogenetic Family Primary Targets	Receptor	Phase 2 Phylogenetic Family Secondary Targets	Receptor
	A19	5-HT_{1B}	A19	5-HT_{1B}
			A19	5-HT_{1D}
			A17	5-HT_{2B} (h)
	A17	5-HT_{2B} (h)	A17	5-HT_{2A} (h)
			A17	5-HT_{2C} (h)
			A19	5-HT_{7} (h)
	A17	5-HT_{6} (h)	▶ A17	5-HT_{6} (h)
	–	5-HT_{4e} (h)	▶ –	5-HT_{4e} (h)
Somatostatin	A04	sst_4 (h)	A04	sst_4 (h)
			A04	sst_5 (h)

Notes: Analysis (Joost and Methner, *Genome Biol.*, 3(11): research 0063.1-0063.16, 2002. http://genomebiology.com/2002/3/11/research/0063). Profiling is performed in two phases. The first phase consists of a panel of 25 GPCR representing the families most commonly associated with off-target effects. Once a compound shows activity in one of the assays from the first set, full dose response is run on receptors within the same target family to assess selectivity (Phase 2).

more than one target can be therapeutic benefit as some disease may require a multitarget therapy.[56] The *in vitro* pharmacological profile of optimized treatment of benign prostate hyperplasia has thus been published and patented.[57] The design of pharmacological profiles aiming to predict adverse effects of drug candidates is a mix of both a rational molecular approach on the basis of phylogenetic clustering of targets and an empirical approach. This has not allowed pharmacologists to come to a consensus on the best predictive profile.

Recently, programs aiming to build a more rational profiling strategy have been launched. These are on the basis of the generation of large datasets consisting of *in vitro* profile results of known drugs in a very broad battery of assays. The statistical analysis of correlations between *in vitro* pharmacological profiles and side effects reported for these drugs allows the linking of *in vitro* data to *in vivo* effects. Cerep has pioneered this approach with the development of Bioprint®, which currently includes around 2500 compounds (including nearly all marketed and recently withdrawn drugs), all tested in over 180 *in vitro* assays.[58,59] Molecules are first tested at one given concentration and a full dose response is determined for each active compound. This dataset gives a clear picture of the frequency of overlap or crosscorrelation between assays. This information is used to design optimized multitiered profiling strategies.

As most of the 2500 compounds have a history of use in humans, a corollary dataset containing human adverse reactions and PK information has been compiled. These datasets are mined to develop quantitative relationships between *in vitro* profiles and *in vivo* effects.

6.11.3.3 Methods

6.11.3.3.1 Quality, Turnaround Time, and Cost

As profiling plays a key role in lead selection and optimization, data quality, availability, and cost become important considerations. Indeed, all profiling data are critical since these will both support the selection of the best candidates and will be included in a corporate database to support further improvement of the discovery process.

TABLE 6.4

Some Examples of Selected Targets and Related Known Information*

Family	Receptor	Source	Therapeutic Area	Tissue Function
Nonpeptide receptors				
Adenosine	A_1	Human cDNA	Cardiac arrhythmia, asthma, myocardial ischemia, obesity, pain, renal disease, sleep apnea, stroke, cancer, inflammation, glaucoma, cystic fibrosis, Alzheimer's disease, Parkinson's disease	Bradycardia, lipolysis inhibition, reduction of glomerular filtration and natriuresis, tubero-glomerular feedback, antinociception, renal vasodilatation-constriction, reduction of central cholinergic and noradrenergic nerve activity, presynaptic inhibition of excitatory neurotransmission
Cannabinoid	CB_1	Human cDNA	Motor function, memory, analgesia, convulsion, Parkinson's disease, emesis, glaucoma, pain, cancer	Modulation of neurotransmitter release, sleep disturbance, weight loss, antiemetic activity, locomotor disfunction, bronchodilatation, neuroprotection, memory loss
GABA	$GABA_A$	Rat brain	Alcoholism, CNS, anesthesia, epilepsy, anxiety, depression, Parkinson's disease, pain, migraine, respiratory, gastrointestinal, diabetes, obesity	Synaptic inhibition, decrease in neuronal excitability, increase in cellular chloride influx, neuroprotection, sedation/hypnosis, anticonvulsant activity, muscle relaxation
Histamine	H_1	Human cDNA	Allergy, inflammation, asthma, chronic sinusitis, rhinitis, epilepsy, migraine, motion sickness, pain, cancer	Smooth muscle contraction, increase in vascular permeability, stimulation of hormone release (adrenal medulla), increase in neuronal firing
Muscarinic	M_1	Human cDNA	Alzheimer's disease, Parkinson's disease, pain, peptic ulcers, inflammation, diabetes	Ganglionic depolarization, vasoconstriction, blood pressure changes, neuroprotection
Purinergic	P2X	Rat urinary bladder	Pain, diabetes	Smooth muscle contraction, nociception
Serotonin	$5\text{-}HT_{1A}$	Human cDNA	Alzheimer's disease, anxiety, depression, schizophrenia, hypertension, inflammation, pain, migraine, spasticity, ulcers, obesity, glaucoma	Somatodendritic autoreceptor in hippocampus and raphe nuclei, circadian rhythm, somatodendritic heteroreceptor at cholinergic terminals of myenteric plexus
Sigma	Nonselective sigma	Rat brain	Inflammation, depression, schizophrenia, amnesia, stroke, cancer, ulcer, pain, diabetes	Regulation of motor function, smooth muscle contraction, neurotransmitter release, neuroprotection

Peptide receptors				
Angiotensin-II	AT_1	Human cDNA	Artherosclerosis, cardiac hypertrophy, congestive heart failure, hypertension, myocardial infarction, renal disease, cancer, diabetes, obesity, glaucoma, cystic fibrosis, Alzheimer's disease, Parkinson's disease	Smooth muscle contraction, cell proliferation and migration, aldosterone and ADH release, central and peripheral sympathetic stimulation, extracellular matrix formation, tubular sodium retention, neuroprotection
Bradykinin	B_2	Human cDNA	Asthma, arthritis, cancer, hypertension, inflammation, migraine, myocardial ischemia, pain, rhinitis, diabetes, cystic fibrosis, nociception	Vasodilatation, stimulation of natriuresis-diuresis in kidney, smooth muscle contraction, induction of hyperalgesia, edema, neuroprotection
Chemokines	CCR1	Human cDNA	Inflammatory disorders, immunosuppression, dermatology, cancer	Monocyte and granulocyte chemotaxis, leukocyte trafficking
Cholecystokinin	CCK_A (CCK_1)	Human cDNA	Anorexia, gastrointestinal disorders, pancreatitis, satiety, diabetes, obesity, nociception, schizophrenia, Parkinson's disease, addiction, cancer	Pancreatic enzyme secretion, potentiation of opiate analgesia, gallbladder contraction, gut motility, growth promoting effects on some tumors, dopamine release
Endothelin	ET_A	Human cDNA	Acute pancreatitis, artherosclerosis, cancer, hypertension, congestive heart failure, diabetes, obesity, inflammation, myocardial ischemia, prostatic hypertrophy, pulmonary fibrosis, stroke, ulcer, pain	Vasoconstriction, bronchoconstriction, positive cardiac inotropy, proliferative responses, aldosterone secretion, neuroprotection
Galanin	GAL1	Human cDNA	Alzheimer's disease, feeding diabetes, growth disorders, pain, stroke, obesity, Parkinson's disease	Inhibition of acetylcholine release, regulation of motility, inhibition of insulin release, stimulation of growth hormone, inhibition of LH-RH secretion
Neurokinin	NK_1	Human cell line	Asthma, hypertension, inflammation, pain, multiple sclerosis, cancer, diabetes, anxiety, social disorders, Alzheimer's disease, Parkinson's disease	Vasorelaxation, bronchodilatation, intestinal smooth muscle contraction, histamine release, plasma protein exudation, neuroprotection
Neuropeptide Y	Y_1	Human cell line	Diabetes, obesity, Alzheimer's disease, anxiety, depression, congestive heart failure, hypertension, ulcers, pain	Vasoconstriction, orexigenic effects, energy homeostasis, neuroprotection
Neurotensin	NT_1 (NTS1)	Human cDNA	Anorexia, thermoregulation, inflammation, pain, Parkinson's disease, schizophrenia, diabetes, Alzheimer's disease	Modulation of dopamine transmission, inhibition of gastric acid secretion, modulation of gastric motility, hypothermia
Vasoactive Intestinal Peptide	VIP_1 ($VPAC_1$)	Human cDNA	Asthma, hypertension, gastrointestinal disorders, inflammation, septic shock, diabetes	Relaxation of smooth muscle, immune modulator, neuroprotection

(*Continued*)

TABLE 6.4 (*Continued*)

Family	Receptor	Source	Therapeutic Area	Tissue Function
Vasopressin	V_{1a}	Human cDNA	Hypertension, congestive heart failure, hyponatremia, coagulation, cardiac arrhythmia, learning, memory, Alzheimer's disease, obesity	Vasoconstriction, cell proliferation, stimulation of platelet aggregation, stimulation of hepatic glycogenesis, steroid secretion, neuroprotection
Nuclear receptors				
Androgen	AR	Human cell line	Cancer, andropause, menopause, Sjogren's syndrome, idiopathic hirsutism, infertility, hypertension, mood disorders, diabetes, cystic fibrosis	Immune modulation, vascular hemodynamics, neuroprotection, skeletal development, reproduction, maintenance of neuronal morphology, regulation of vasopressin V_{1a} expression, prostatic blood flow and LH production
Ion channels				
K^+ channels	K_{ATP}	Rat brain	Diabetes, asthma, cardiac arrythmia, angina, cardiac and cerebral ischemia, thrombosis, hypertension, incontinence, pain, neurogenic inflammation, epilepsy, stroke, hair growth	Control of insulin release, vasodilatation, protection against cell ischemia
Na^+ channel	Site 2	Rat brain	Cardiac arrhythmia, epilepsy, Alzheimer's disease, pain, gastrointestinal, stroke, glaucoma	Control of neuronal and cardiac electrical activity, neuroprotection
Amine transporters				
Norepinephrine	NE transporter	Human cDNA	Depression, Alzheimer's disease, epilepsy, anxiety, attention deficit hyperactivity, angina, asthma, cardiac arrhythmia, cardiac hypertrophy, congestive heart failure, myocardial ischemia, hypertension, artherosclerosis, narcolepsy, orthostatic hypotension, prostatic hyperplasia, rhinitis, diabetes, diarrhea, glaucoma, impotence, obesity, opiate withdrawal pain, Raynaud's disease, preterm labor pain	Modulation of norepinephrine concentration in the neuronal synaptic clefts, neuroprotection
Enzymes				
Kinases	Abl kinase	Mouse cDNA	Cancer, acute lymphocytic leukemia, chronic myelocytic leukemia, chronic neutrophilic leukemia	Regulation of cytoskeletal organization and movement
Proteases	Elastase	Human leukocytes	Pain, inflammation, cancer, dermatology, cystic fibrosis, immunosuppression	Induction of atherosclerosis, acute hemorrhagic pancreatitis
	MMP1	Human cDNA	Pain, inflammation, cancer, glaucoma, diabetes, dermatology, immunosuppression, angiogenesis	Normal and pathologic tissue remodeling process, wound healing, angiogenesis, tumor invasion
Phosphodiesterases	PDE2	Human cell line	Inflammation, congestive heart failure, arrythmias, obesity	Cardiac positive inotropy and chronotropy

*A complete list of target–function relationships is available at Cerep [URL: http://www.cerep.fr/] on request.

Data validation should follow a clear, well-described and documented process. First, *in vitro* models must be fully validated and robust to give reliable and reproducible data. Experiments then have to be conducted with very strict quality standards. All data must be generated under similar experimental procedures using similar reagents. Assay validation criteria have to be well defined. This includes high signal-to-noise ratio, reproducible inhibitory concentration resulting in a 50% reduction in activity (IC_{50}) or effective concentration resulting in a 50% stimulation in activity (EC_{50}) values of reference compounds, Hill number (n_H), Z' value to assess robustness of the assay,[60] homogeneity of replicates (standard deviation $<20\%$) and reproducibility from one experiment to another.

Moreover, compound solubilization and potential interference of the compounds with the detection methods (quenching, nonspecific binding, etc.) must be systematically characterized.

Profiling has benefited from the recent technical improvements in HTS technologies, most importantly in assay miniaturization and automation.[61,62] Almost all assays currently used in profiling are reliable, cost-effective, and have high throughput. These technological improvements, linked to a reduction in cost of data generation and evidence for the usefulness of the generated information, have contributed to make profiling a key step in drug discovery.

6.11.3.3.2 Binding vs. Functional Assays

The principal aim of the profile is to characterize compounds using simple systems able to predict *in vivo* effect. *In vitro* receptor-binding assays are generally included in profiles to assess off-target interactions. However, when a compound is active in a receptor-binding assay, it is not clear if it will act as an antagonist or agonist. Agonist or antagonist activities can only be displayed through the use of more complex systems:

- *In vivo* assays
- *In vitro* assays using isolated organs
- Whole-cell assays (allowing measurement of phosphoinositide hydrolysis, adenylate cyclase activity, etc.)
- Subcellular assays (G-protein activity).

Functional assays provide enriched information since in a single assay both affinity and likely biological effect are determined. However, to be included in large profile, each assay must meet quality, turnaround time, and cost constraints. Complex whole-cell or -tissue assays may make the interpretation of the data difficult, as the measured effect may not be result of the interaction of the compound tested with the primary target (many targets are present). For this reason, the use of *in vitro* functional assays is generally considered in secondary assays, i.e., assays that aim to determine the functional effect of a compound once it is shown to bind the target of interest. Thus, *in vitro* receptor binding and isolated enzyme assays remain the assays of choice in early profiling.

6.12 Predicting Bioavailability: Pharmaceutical Profiling to Assess Absorption, Distribution, Metabolism, and Excretion Properties

Bioavailability can be assessed early in the discovery process by combining data from a number of independent *in vitro* assays run as part of a pharmaceutical properties profile. These assays generally include some measurement of aqueous solubility, lipophilicity, cell or membrane permeability, and metabolic stability. In most cases, proper interpretation of

data from a pharmaceutical properties profile will give reliable guidance for the identification of compounds with reasonable bioavailability. When testing large numbers of diverse compounds a set of categorical rules on the basis of the knowledge derived from, or validated by, profiling data from large numbers of reference compounds with known *in vivo* properties can be used for guidance. Examples for some commonly measured parameters and values associated with moderate to good bioavailability are listed in Table 6.5. It should be noted that there are examples of successful drugs falling outside of these guidelines, and information on the likely route of administration, dosage, formulation, serum protein binding, ionization effects, etc., need to be considered in addition to standard results from a pharmaceutical profile.

To bring a relevant context to pharmaceutical profile interpretation, it is important to understand how currently marketed drugs behave in each assay. Results from the profiling of approximately 300 marketed drugs with well-documented human % absorption and bioavailability are shown in Figure 6.7 and Figure 6.8. In this example, solubility was measured by adding dimethyl sulfoxide (DMSO)-solubilized drug to an isotonic pH 7.4 aqueous buffer to yield a final concentration of 2% DMSO and a theoretical 200 μM test compound. Ultrafiltration is used to remove precipitated drug and final concentration is measured by liquid chromatography–ultraviolet (LC–UV) or liquid chromatography–mass spectrometry detection (LC–MS). Lipophilicity is measured as by *n*-octanol/PBS pH 7.4 partitioning (log *D*). Apparent permeability was measured using a subclone of the Caco2 cell line. Metabolic stability was measured by incubating drug at 1 μM for 60 min with human liver microsomes. In Figure 6.7 and Figure 6.8, a red flag is given to a compound having solubility less than 20 μM, log $D<-1$ or >4, apparent permeability less than 5×10^{-6} cm/sec, or metabolic stability less than 40% remaining.

A well-designed apparent permeability assay can give guidance in the identification of compounds with desired human absorption (Figure 6.7). The low contributions made by solubility and lipophilicity data alone are not surprising as compounds with poor solubility or extreme log *D* values are likely to also show low apparent permeability. The situation with bioavailability is more complex (Figure 6.8). In this case, there are the expected independent contributions to human bioavailability of both *in vitro* permeability and metabolic stability assay results. In the case of bioavailability, the most significant guidance is given by the "one or more red flag" results. It is important not to over interpret individual assay results. Note that nearly 30% of the drugs having less than 20% oral bioavailability did not have any red flag for solubility, log *D* values, permeability, or metabolic stability. At the other end of the spectrum, ~15% of the compounds with 80 to 100% oral bioavailability did have one or more red flags (Figure 6.8).

To further strengthen decision-making, it is helpful to use statistical clustering to place each experimental compound in the context of a database created from the pharmaceutical profiling of compounds with well-characterized human PK. In the absence of

TABLE 6.5

Bioavailability Assessment through the Use of *In Vitro* Assays and Values Associated with Moderate to Good Bioavailability

Pharmaceutical Properties	**Quantified Results Obtained in Corresponding *In Vitro* Assay**
Aqueous solubility (equilibrium)	>20 μM
Lipophilicity	Log *D* between −1 and 4
Apparent permeability	$> 5 \times 10^{-6}$ cm sec $^{-1}$
Metabolic stability	> 40% remaining

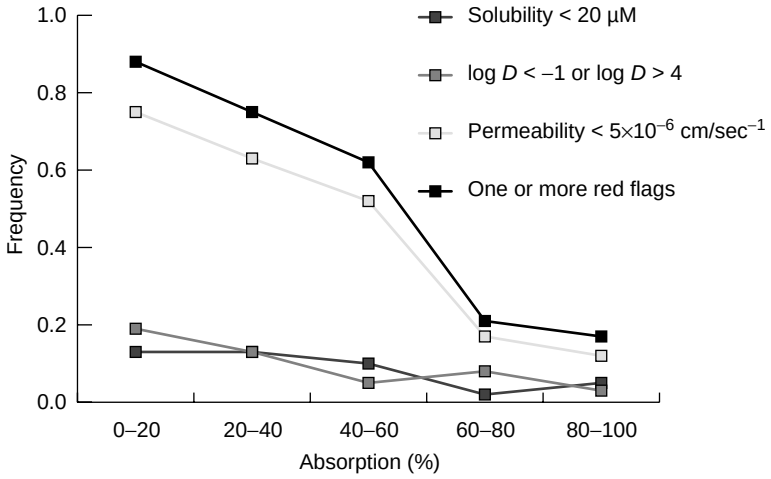


FIGURE 6.7
Approximately 300 drugs were tested in aqueous solubility, log *D*, and apparent permeability assays. Compounds having low values for solubility, or apparent permeability, or extreme log *D* values were flagged. The frequency of compounds with flags in each human absorption (%) bin is shown.

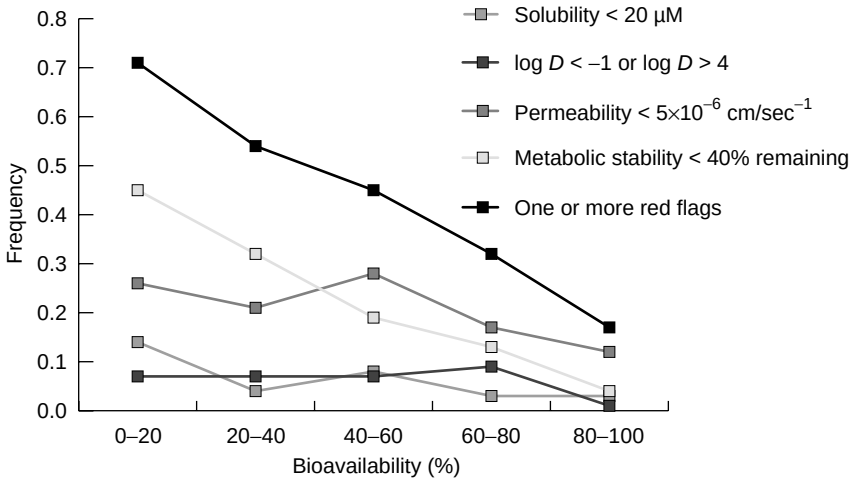


FIGURE 6.8
Approximately 300 drugs were tested in aqueous solubility, log *D*, apparent permeability and metabolic stability assays. Compounds having low values for solubility, apparent permeability, or metabolic stability, or extreme log *D* values were flagged. The frequency of compounds with flags in each human bioavailability (%) bin is shown.

additional guiding information, it is reasonable to assume that compounds with overall similar pharmaceutical profiles will have similar *in vivo* PK properties. Table 6.6 shows a single test compound that clusters with several marketed drugs on the basis of the similarity of *in vitro* profile results. Human absorption and oral bioavailability are shown for the drugs. From this table it can be seen that the group of selected drugs with favorable apparent permeability, solubility, and log *D* values all have good human absorption. It is reasonable to project that the test compound will be likely to have similar *in vivo* absorption. However, there is a distribution of 22 to 69% remaining for the *in vitro*

metabolic stability assay. As all compounds shown have good apparent permeability it would be appropriate to focus attention on the compounds with metabolic stability most similar to the test compound. A reasonable assumption would be that the test compound should be flagged as having a potentially low bioavailability (15 to 20%) (Table 6.6).

Additional assays used in early pharmaceutical property profiles usually include plasma protein binding, individual cytochrome P450 assays, stability in the presence of serum, production of metabolites likely to be involved in covalent binding to biomolecules, and interaction with efflux pumps:

- Serum protein binding is an important contributor to distribution and clearance *in vivo*. In addition, results from *in vitro* tests of compounds with high serum protein binding are expected to be sensitive to the presence of serum or other proteins in assay incubations.
- Panels of cytochrome P450 assays are run to identify compounds that interact as substrates or inhibitors of the major drug-metabolizing enzymes. P450 results in early drug discovery are often used to guide medicinal chemists away from compounds that interact primarily with the P450s that are polymorphic in the target patient population.
- Serum stability is tested when it is suspected or desired that serum enzymes may hydrolyze the parent drug. Some compounds or their metabolites are subject to glutathione conjugation, which may lead to covalent binding to protein and deoxyribonucleic acid (DNA). The potential for glutathione conjugation can be tested *in vitro* using human liver preparations.
- There are several efflux pumps which may affect absorption, blood–brain-barrier penetration, and reabsorption from kidney microtubules. The most commonly tested efflux pump in early drug discovery is the P-glycoprotein. Assays to identify P-glycoprotein substrates or inhibitors can be run using a variety of cell lines.

A variety of *in vitro* assays are available to assess important parameters that impact bioavailability and these pharmaceutical profiling assays are commonly run to support drug discovery efforts. The major challenge is the interpretation of these results. Misinterpretation can lead to dropping promising compounds or alternatively taking

TABLE 6.6

Single Test Compound Clustering with Several Marketed Drugs on the Basis of the Similarity of *In Vitro* Profile Results Including Human Absorption, Oral Bioavailability, Permeability, Solubility, Log *D*, and Metabolic Stability Characteristics

Compound	Solubility (μM)	Log *D*	Apparent Permeability (10^{-6} cm/sec)	Metabolic Stability (% Remaining)	Oral Bioavailability (%)	Absorption (%)
Test compound	123	3.9	36	26	N/A	N/A
Drug A	210	3.3	117	22	15	93
Drug B	180	3.1	17	27	23	80
Drug C	128	3.6	32	51	15	N/A
Drug D	205	3.8	55	69	18	53
Drug E	113	3.3	25	68	75	75
Drug F	23	3.6	29	20	17	43
Drug G	33	3.5	33	28	60	60
Drug H	79	2.9	51	30	15	95

unsuitable compounds into early development. In our experience, the interpretation of pharmaceutical profiling data is greatly strengthened by understanding how the active ingredients of all currently marketed drugs behave in these same *in vitro* assays.

6.13 Position of Profiling in Current Drug Discovery Process: When Profiling Should Be Performed

Originally, profiling was used as a final step in the characterization of advanced lead compounds. However, as its added value is increasingly recognized, it has been moved to early in the discovery process. This evolution has been aided by improvements in screening technologies that yield reduction of cost and time in the generation of large profiles. Today, the question of whether drug profiling must occur just after hit identification or later in the lead optimization process is often debated.

6.13.1 The Sequential vs. the Parallel Approach

Drug discovery was initially a sequential process and the first priority was the optimization of affinity for the therapeutic target. Each lead optimization cycle would then generate more active compounds, independently of any specific criteria. Once an acceptable potency was achieved, leads were then tested for selectivity and specificity against a large number of biological targets ("lead profiling"). This conventional drug discovery on the basis of a sequential rationale is described in Figure 6.9. It generally involves, in succession:

- Evaluation of affinity of many compounds for a therapeutic target
- Chemical optimization of affinity of the most active compounds

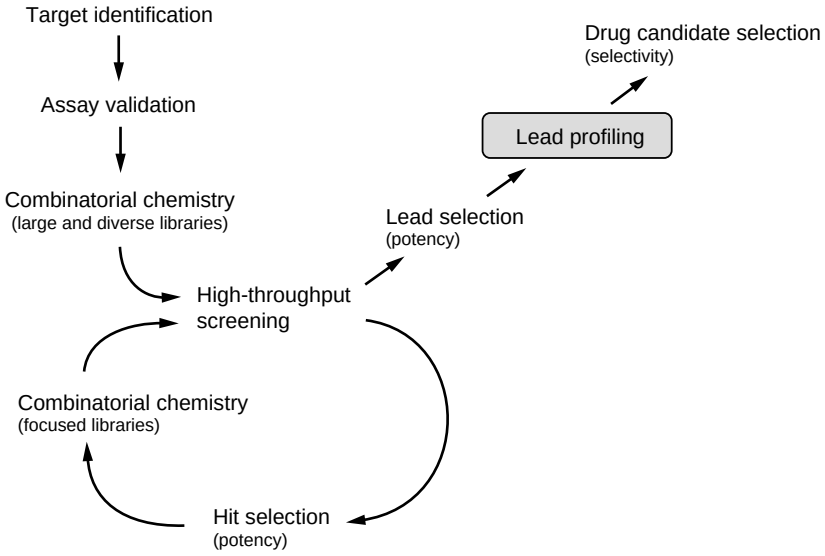


FIGURE 6.9
Sequential drug discovery process and lead profiling.

- Chemical optimization of the selectivity and specificity of the most potent compounds
- Chemical optimization of the bioavailability and safety of the most promising compounds.

The success of this method relies on the hypothesis that the chemical optimization of the molecule for its affinity for one target will not increase its affinity for unwanted targets. There are, however, many examples of discovery programs where hypothesis did not hold.

Changes in the drug discovery process have led to the emergence of a parallel approach, where the characteristics of the compound (potency, specificity, selectivity, and bioavailability) are taken into account early in the process. Hits are profiled just after the HTS step (Figure 6.10). Lead optimization then becomes a multidimensional approach, where both pharmacological and pharmaceutical characteristics are balanced and optimized simultaneously.[63] Lead optimization becomes more challenging for chemists since they will have to optimize for more than one characteristic simultaneously. However, there is evidence that close analogs of "clean" compounds (i.e., good selectivity, specificity, and drug-like pharmaceutical profile) are themselves more likely than the average compound to be "clean." [64]

6.13.2 The Reverse Sequential Approach Profiling before Screening

As described above, large databases of both pharmacological and pharmaceutical properties of compounds are now used to support library design. Some pharmaceutical companies have initiated intensive profiling programs to define chemical space that is most suitable for the use of discovering lead compounds against important targets classes.[58] Thus in each new program, libraries of compounds likely to have affinity for the target of interest are designed *in silico*. In such a strategic process, profiling is used prior to screening, and therapeutic target screening is performed on a limited number

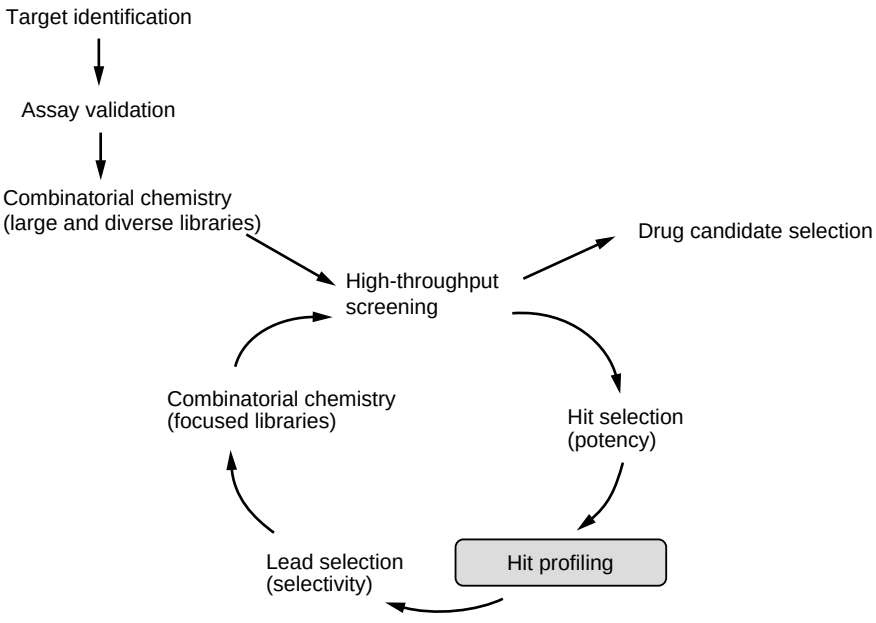


FIGURE 6.10
Parallel drug discovery process and hit profiling.

of compounds having suitable properties for later development. This reverse sequential drug discovery process is described in Figure 6.11. Applying an intensive profiling strategy requires the ability to test large series of molecules in large panels of assays at low cost and with fast turnaround time. Such technologies are commonly referred to as high-throughput profiling.[65]

6.14 Conclusion

Profiling is emerging as a key step in rational discovery. It contributes to reduced attrition in clinical development and thus potentially to increasing the number of new drugs reaching the market. Pharmacological and pharmaceutical profiling, benefiting from the recent advances in molecular biology and from improvement of screening technologies, allow the creation of large databases. Through early and complete understanding of off-target effects and pharmaceutical properties, compounds likely to succeed in development are selected.

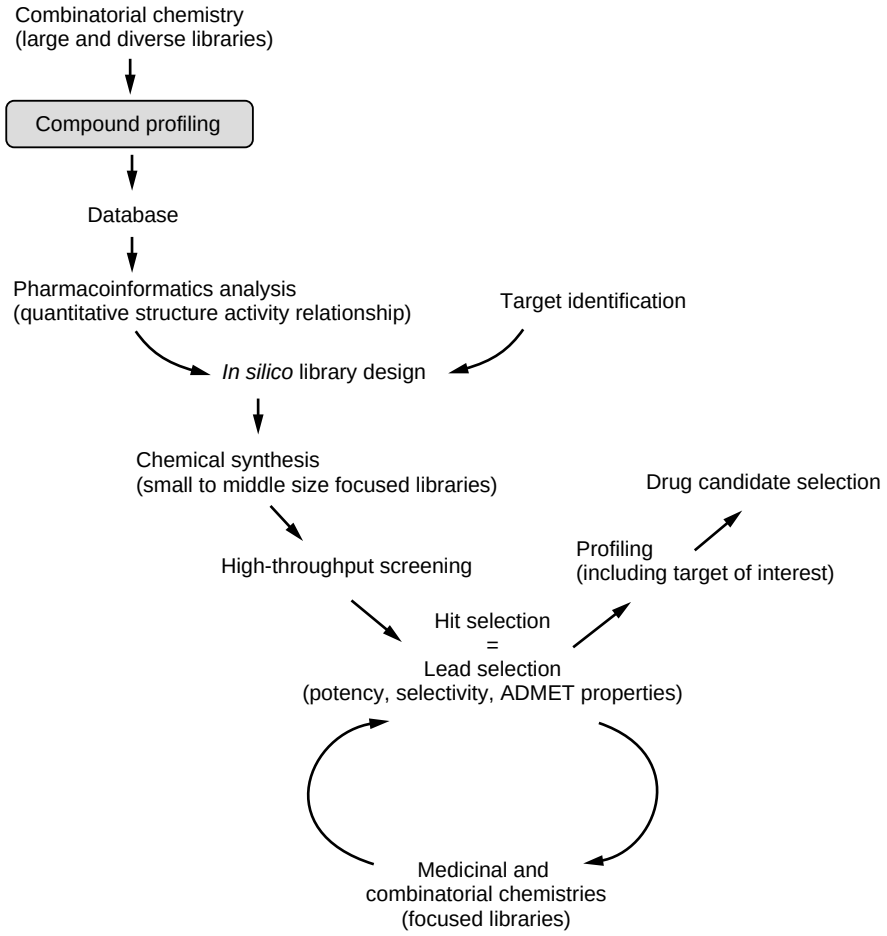


FIGURE 6.11
Reverse sequential drug discovery process and compound profiling.

From a long-term perspective, broad profiling is the basis for building a new understanding of the relationships between small molecules and protein targets. These fundamental relationships will fuel the further advancement of drug discovery and development.

Acknowledgments

The authors thank Cheryl Wu for her contribution to the pharmaceutical profile part; Marie-Christine Bodinier, Emmanuelle Guillaume, Philippe Dupuis, and Gervais Néliat for their contributions to Table 6.4; Catherine Moreau, Cécile Diet, and Loïc Dorgeret for technical assistance.

References

1. Kennedy, T., Managing the drug discovery/development interface, *Drug Discov. Today*, 2, 436, 1997.
2. Pirmohamed, M., James, S., Meakin, S., Green, C., Scott, A.K., Walley, T.J., Farrar, K., Park, B.K., and Breckenridge, A.M., Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients, *Br. Med. J.*, 329, 15, 2004.
3. Einarson, T.R., Drug-related hospital admissions, *Ann. Pharmacother.*, 27, 832, 1993.
4. Anonymous, *ICH S7A: Safety pharmacology studies for human pharmaceuticals*, CPMP/ICH/539/00, 2000.
5. *The GLP Pocket Book: the Good Laboratory Practice Regulations 1999 and Guide to UK GLP Regulations 1999*, MCA Publications, London, 1999.
6. Izquierdo, I., Merlos, M., and Garcia-Rafanell, J., Rupatadine: a new selective histamine H_1 receptor and platelet-activating factor (PAF) antagonist. A review of pharmacological profile and clinical management of allergic rhinitis, *Drugs Today*, 39, 451, 2003.
7. Dougall, I.G., Young, A., Ince, F., and Jackson, D.M., Dual dopamine D_2 receptor and β_2-adrenoceptor agonists for the treatment of chronic obstructive pulmonary disease: the preclinical rationale, *Resp. Med.*, 97, S3, 2003.
8. Ind, P.W., Laitinen, L., Laursen, L., Wenzel, S., Wouters, E., Deamer, L., and Nystrom, P., Early clinical investigation of Viozan™ (sibenadet HCI), a novel D_2 dopamine receptor, β_2-adrenoceptor agonist for the treatment of chronic obstructive pulmonary disease symptoms, *Resp. Med.*, 97, S9, 2003.
9. Holm, K.J. and Markham, A., Mirtazapine: a review of its use in major depression, *Drugs*, 57, 607, 1999.
10. Doogan, D.P., Toleration and safety of sertraline: experience worldwide, *Int. Clin. Psychopharm.*, 6, 47, 1991.
11. Ferguson, J.M., Antidepressant medications: adverse effects and tolerability, *Prim. Care Compan. J. Clin. Psych.*, 3, 22, 1997.
12. Goldstein, I., Lue, T.F., Padma-Nathan, H., Rosen, R.C., Steers, W.D., and Wicker, P.A., Oral sildenafil in the treatment of erectile dysfunction, *New Engl. J. Med.*, 338, 1397, 1998.
13. Curran, M.P. and Keating, G.M., Tadalafil, *Drugs*, 63, 2203, 2003.
14. Laties, A.M., Ellis, P., and Mollon, J.D., The effects of sildenafil citrate (Viagra®) on color discrimination in volunteers and patients with erectile dysfunction, *Inv. Ophthalmol. Vis. Sci.*, 40, S693, 1999.
15. Wallis, R.M., Casey, J., and Howe, L., Characterisation of retinal phosphodiesterase (PDE) isozymes and the effects of sildenafil in vitro, *Ophthalmic Res.*, 30, 111, 1998.
16. Neubig, R.R., Spedding, M., Kenakin, T., and Christopoulos, A., International Union of Pharmacology Committee on Receptor Nomenclature and Drug Classification. XXXVIII. Update on terms and symbols in quantitative pharmacology, *Pharmacol. Rev.*, 55, 597, 2003.

17. Cheng, Y. and Prusoff, W.H., Relationship between the inhibition constant (K1) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction, *Biochem. Pharmacol.*, 22, 3099, 1973.
18. Schild, H.O., pA, a new scale for the measurement of drug antagonism, *Br. J. Pharmacol.*, 2, 189, 1947.
19. Chambers, C., Smith, F., Williams, C., Marcos, S., Liu, Z.H., Hayter, P., Ciaramella, G., Keighley, W., Gribbon, P., and Sewing, A., Measuring intracellular calcium fluxes in high throughput mode, *Comb. Chem. High Throughput Scr.*, 6, 355, 2003.
20. Sullivan, E., Tucker, E.M., and Dale, I.L., Measurement of [Ca^{2+}] using the Fluorometric Imaging Plate Reader (FLIPR), *Methods Mol. Biol.*, 114, 125, 1999.
21. Kenakin, T.P., The classification of drugs and drug receptors in isolated tissues, *Pharmacol. Rev.*, 36, 165, 1984.
22. Borman, R.A., Tilford, N.S., Harmer, D.W., Day, N., Ellis, E.S., Sheldrick, R.L., Carey, J., Coleman, R.A., and Baxter, G.S., 5-HT_{2B} receptors play a key role in mediating the excitatory effects of 5-HT in human colon in vitro, *Br. J. Pharmacol.*, 135, 1144, 2002.
23. Brown, C.M., MacKinnon, A.C., Redfern, W.S., Hicks, P.E., Kilpatrick, A.T., Small, C., Ramcharan, M., Clague, R.U., Clark, R.D., and MacFarlane, C.B., The pharmacology of RS-15385-197, a potent and selective alpha 2-adrenoceptor antagonist, *Br. J. Pharmacol.*, 108, 516, 1993.
24. Heaslip, R.J. and Evans, D.Y., Emetic, central nervous system and pulmonary activities of rolipram in the dog, *Eur. J. Pharmacol.*, 286, 281, 1995.
25. Horowski, R. and SastreHernandez, Y.M., Clinical effects of the neurotropic selective cAMP phosphodiesterase inhibitor rolipram in depressed patients: global evaluation of the preliminary reports, *Curr. Ther. Res. Clin. Exp.*, 38, 23, 1985.
26. Chiba, T., Bharucha, A.E., Thomforde, G.M., Kost, L.J., and Phillips, S.F., Model of rapid gastrointestinal transit in dogs: effects of muscarinic antagonists and a nitric oxide synthase inhibitor, *Neurogastroenterol. Motil.*, 14, 535, 2002.
27. Kay, G.G., Berman, B., Mockoviak, S.H., Morris, C.E., Reeves, D., Starbuck,V., Sukenik, E., and Harris, A.G., Initial and steady-state effects of diphenhydramine and loratadine on sedation, cognition, mood, and psychomotor performance, *Arch. Int. Med.*, 157, 2350, 1997.
28. Motomura, S., Zerkowski, H.R., Daul, A., and Brodde, O.E., On the physiologic role of β_2 adrenoceptors in the human heart: in vitro and in vivo studies, *Am. Heart J.*, 119, 608, 1990.
29. Lewis, S.C., Langman, M.J.S., Laporte, J.R., Matthews, J.N.S., Rawlins, M.D., and Wiholm, B.E., Dose-response relationships between individual nonaspirin nonsteroidal anti-inflammatory drugs (NANSAIDs) and serious upper gastrointestinal bleeding: a meta-analysis based on individual patient data, *Br. J. Clin. Pharmacol.*, 54, 320, 2002.
30. Bran, S., Murray, W.A., Hirsch, I.B., and Palmer, J.P., Long QT syndrome during high-dose cisapride, *Arch. Int. Med.*, 155, 765, 1995.
31. Skold, A.C., Danielsson, C., Linder, B., and Danielsson, B.R., Teratogenicity of the IKr-blocker cisapride: relation to embryonic cardiac arrhythmia, *Reprod. Toxicol.*, 16, 333, 2002.
32. Redfern, W.S., Carlsson, L., Davis, A.S., Lynch, W.G., MacKenzie, I., Palethorpe, S., Siegl, P.K., Strang, I., Sullivan, A.T., Wallis, R., Camm, A.J., and Hammond, T.G., Relationships between preclinical cardiac electrophysiology, clinical QT interval prolongation and torsade de pointes for a broad range of drugs: evidence for a provisional safety margin in drug development, *Cardiovasc. Res.*, 58, 32, 2003.
33. Dogterom, P., Zbinden, G., and Reznik, G.K., Cardiotoxicity of vasodilators and positive inotropic/vasodilating drugs in dogs: an overview, *Crit. Rev. Toxicol.*, 22, 203, 1992.
34. Herman, E.H., Ferrans, V.J., Young, R.S., and Balazs, T., A comparative study of minoxidil-induced myocardial lesions in beagle dogs and miniature swine, *Toxicol. Pathol.*, 17, 182, 1989.
35. Sobota, J.T., Review of cardiovascular findings in humans treated with minoxidil, *Toxicol. Pathol.*, 17, 193, 1989.
36. Beld, A.J., Van den Hoven, S., Wouterse, A.C., and Zegers, M.A.P., Are muscarinic receptors in the central and peripheral nervous system different?, *Eur. J. Pharmacol.*, 30, 360, 1975.
37. Holck, M.I., Marks, B.H., and Wilberding, C.A., Characterization of alpha-adrenergic receptors in guinea-pig vas deferens by [^{3}H]dihydroergocryptine binding, *Mol. Pharmacol.*, 16, 77, 1979.
38. Leclerc, G., Rouot, B., Velly, J., and Schwartz, J., β-Adrenergic receptor subtypes, *Trends Pharmacol. Sci.*, 2, 18, 1981.

39. Nademanee, K. and Singh, B.N., Control of cardiac arrhythmias by calcium antagonism, *Ann. NY Acad. Sci.*, 522, 536, 1988.
40. Ekins, S., Predicting undesirable drug interactions with promiscuous proteins *in silico*, *Drug Discov. Today*, 9, 276, 2004.
41. Altman, R.B. and Klein, T.E., Challenges for biomedical informatics and pharmacogenomics, *Annu. Rev. Pharmacol. Toxicol.*, 42, 113, 2002.
42. Kola, I. and Landis, J., Can the pharmaceutical industry reduce attrition rates? *Nat. Rev. Drug Discov.*, 3, 711, 2004.
43. Franz, S., 2003 approvals: a year of innovation and upward trends, *Nat. Rev. Drug Discov.*, 3, 103, 2004.
44. Centre for Medicines Research: http://www.cmr.org, Innovation on the wave? CMR International, 2004.
45. Dickson, M. and Gagnon, J.P., Key factors in the rising cost of new drug discovery and development, *Nat. Rev. Drug Discov.*, 3, 417, 2004.
46. Centre for Medicines Research: http://www.cmr.org, R&D Pharmaceutical investment and output survey, CMR International, 2003.
47. Booth, B. and Zemmel, R., Prospects for productivity, *Nat. Rev. Drug Discov.*, 3, 451, 2004.
48. Kola, I. and Rafferty, M., New technologies that may impact drug discovery in the 5-10 year timeframe workshop, 2002 Biomed. Expo Ann Arbor Michigan, Data derived from Prous Science, Drug News Prospect, 2002.
49. DiMasi, J.A., Hansen, R.W., and Grabowski, H.G., The price of innovation: new estimates of drug development costs, *J. Health Econom.*, 22, 151, 2003.
50. Drews, J., Drug discovery: a historical perspective, *Science*, 287, 1960, 2000.
51. Kennedy, T., Managing the drug discovery/development interface, *Drug Discov. Today*, 2, 436, 1997.
52. Prentis, R.A., Lis, Y., and Walker, S.R., Pharmaceutical innovation by the seven UK-owned pharmaceutical companies (1964–1985), *Br. J. Clin. Pharmacol.*, 25, 387, 1988.
53. Burrill, G.S., Fewer drugs approved, more money spent. Where's the beef? *Drug Discov. World*, 4, 9, 2004.
54. Joost, P. and Methner, A., Phylogenic analysis of 277 human G-protein-coupled receptors as a tool for the prediction of orphan receptor ligands, *Genome Biol.*, 3: research 0063.1-0063.16, 2002. http://genomebiology.com/2002/3/11/research/0063
55. Curran, M.E., Splawski, I., Timothy, K.W., Vincent, G.M., Green, E.D., and Keating, M.T., A molecular basis for cardiac arrhythmia: HERG mutations cause long QT syndrome, *Cell*, 80, 795, 1995.
56. Morphy, R., Kay, C., and Rankovic, Z., From magic bullets to designed multiple ligands, *Drug Discov. Today*, 9, 641, 2004.
57. Gluchowski, C., Forray, C.C., Chui, G., Branchek, T.A., Weltzel, J.M., and Harting, P.R., Use of Alpha 1c Specific Compounds to Treat Benign Prostatic Hyperplasia, Patent number 5,403,847, Synaptic Pharmaceutical, 1995.
58. Fliri, A.F., Loging, W.T., Thadeio, P.F., and Volkmann, R.A., Biological spectra analysis: linking biological activity profiles to molecular structure, *PNAS*, 102, 261, 2005.
59. Krejsa, C.M., Horvath, D., Rogalski, S.L., Penzotti, J.E., Mao, B., Barbosa, F., and Migeon, J.C., Predicting ADME properties and side effects: the BioPrint approach, *Curr. Opin. Drug Discov. Dev.*, 6, 470, 2003.
60. Zhang, J., Chung, T.D.Y., and Oldenburg, K.R., A simple statistical parameter for use in evaluation and validation of high throughput screening assays, *J. Biomol. Screen.*, 4, 67, 1999.
61. Valler, M.J. and Green, D., Diversity screening versus focused screening in drug discovery, *Drug Discov. Today*, 5, 286, 2000.
62. Walters, W.P. and Namchuk, M., Designing screens: how to make your hits a hit, *Nat. Rev. Drug Discov.*, 2, 259, 2003.
63. Bleicher, K.H., Böhm, H.-J., Müller, K., and Alanine, A.I., Hit and lead generation: beyond high-throughput screening, *Nat. Rev. Drug Discov.*, 2, 369, 2003.
64. Poulain, R., Horvath, D., Bonnet, B., Eckhoff, C., Chapelain, B., Bodinier, M.-C., and Deprez, B., From hit to lead. Combining two complementary methods for focused library design. Application to μ opiate ligands, *J. Med. Chem.*, 44, 3378, 2001.
65. Jean, T. and Chapelain, B., International publication number WO-09915894, Method of identification of leads or active compounds, CEREP, 1999.

7

Cell-Based Analysis of Drug Response Using Moving Optical Gradient Fields

Jeff M. Hall, Ilona Kariv, Patricia A. McNeeley, Phillip J. Marchand, and Tina S. Nova

CONTENTS

7.1 Introduction

In vitro testing platforms for the determination and monitoring of the actions of a drug can be used in four phases of drug discovery, development, and application:

- In the preclinical discovery of new drugs and their biochemical actions on the cell
- In the preclinical development and optimization of an active compound
- As a surrogate or clinical marker of efficacy during all phases of human clinical trials
- For marketed compounds as a pharmacogenomic measure, predicting which patients in a broader population will best respond to the drug and, therefore, who should be prescribed a given drug.

Many technology platforms have utility for several of these stages but, rarely, can the same platform be used across all four stages. An ideal platform would be one that could be used in preclinical work on cell lines and primary tissue samples and, then, also be used in the clinical phase on tissue samples from trial subjects and, finally, could be applied to biological samples obtained from patients in the general population who might be candidates for a given drug. If there were a predictive value to such a test it would be optimal if it could be used for determining the best dosage of a drug for a particular patient.

Cell-based assays address the need for biologically relevant data from all four phases of screening listed above. Cell-based screening can provide information on multiple parameters for a given target or multiple targets simultaneously in a biologically relevant context. Increasing the quality of output from primary screening is particularly important given the current influx of new nonvalidated drug targets emerging from genomics with which the pharmaceutical industry has little or no past experience. New technologies that can provide reliable detection of cell-based assay systems are required to meet the needs described above. Miniaturization of such systems is also required as this will minimize the need for expensive, large quantities of cell lines or primary human tissue. Constructing assays that are predictive of how inter-individual clinical responses to drugs is a challenge, which is only now being met by the novel technologies that are under development.

Optophoresis™ is one such technology. It has been developed to provide analysis of the drug response of small numbers of cells without any need for labeling or cell processing. *Optophoresis* is defined as the analysis of the motion of cells, where the motion is either induced or modified by a moving optical gradient field, which produces radiation pressure forces on the cells in an aqueous suspension. Quantitation of the induced motion provides a basis for distinguishing one population of cells, such as those responding to a drug, from another population, such as cells that are not responding to the same drug. The optophoretic fast-scan technique measures the distribution of distances traversed by a population of cells when exposed to a fast-moving optical gradient. It does so by imaging and measuring the distance of each individual cell in a field of view. The fast-scan measurement technique has been validated using cell-based apoptosis models in leukemia cell lines.[1] Optophoretic measurements have been quantitatively comparable to reference assays with regard to drug selectivity and potency and target specificity, demonstrating the suitability of this technology for pharmaceutical and clinical diagnostic applications.[2]

Optophoresis has the advantage over other cell-based assays in that a small population of cells is used, and individual data on the drug response of each cell in the population is acquired. In this way subpopulations in a heterogeneous sample can be detected and analyzed. In addition, because of the low number of cells required, the cells do not need to be grown out or manipulated in any way. They are merely incubated with the drug *ex vivo* for the appropriate time to obtain a cellular response and then measured.

7.2 Measurement Technique and Instrument

Optophoresis builds on prior demonstrations that light carries momentum and that this momentum may be transferred to a material particle, resulting in a radiation pressure force on that particle. Ashkin[3] was the first to unambiguously demonstrate radiation pressure forces by applying the intense optical field from a collimated laser beam to nearly transparent, micron-scale particles in suspension. The ability to create intense, optical gradients using lasers eventually led to many new technologies, such as laser tweezers. Laser tweezers rely on a highly focused laser beam to create a three-dimensional optical trap that enables capture and manipulation of individual particles,[4,5] including cells,[6] viruses, and bacteria[7] (for review see Svoboda and Block[8]). These and similar trapping technologies have been used to measure the forces involved in molecular binding,[9,10] in the motility of sperm and bacteria,[11,12] and in the measurement of the resistance of erythrocytes to deformation by external forces.[13,14] In none of these applications has there been an attempt to measure quantitatively the pharmacological response of cells. Optophoresis provides an ability to evaluate cellular responses to a drug treatment inferred from quantitative measurements of the

transverse motion of a cell induced by a moving optical gradient. This gradient produces forces on an illuminated particle suspended in a medium due to a spatially nonuniform optical intensity field, such as the Gaussian intensity profile that is characteristic of a laser beam.

Because cells are suspended in an aqueous medium, they are subject to forces other than optical, principally viscous drag, surface interactions, and gravity. The relative strengths of all the forces depend on physical properties of the various constituents of the system. The strengths of the transverse and axial optical forces depend on the indices of refraction of the cell and the medium, the cell's size and morphology, and the intensity, shape, and wavelength of the optical gradient. Gravity, corrected for the buoyancy of the medium, depends on the cell's volume and the respective densities of the cell and the medium. Viscous drag is a result of the cell's motion in the medium and is determined by the viscosity of the medium and the cell's size, morphology, and velocity. Forces created by interactions between the cell and nearby surfaces are dependent on the substrate surface properties and the cell's size, morphology, and membrane composition. Under the influence of gravity, the cells settle on the bottom substrate of the sample well. When a cell is exposed to a moving optical gradient, the optical force drives the motion of the cell. The axial component typically opposes the gravitational force, whereas the transverse component pulls the cell toward the maximum of the optical intensity. Viscous drag and surface interactions produce dissipative forces that oppose the motion of the cell. The sum of the optical force (which depends on the size, morphology, and internal structure of the cell) and the dissipative forces (reflecting the size, shape, and morphology of the cell and its membrane) determine the magnitude of a cell's motion when it is exposed to a moving optical gradient.

Biological differences between two cell populations can be assayed with the moving optical gradient by various methods. For example, the speed of translation of the optical gradient at which the dissipative forces of viscous drag and surface interaction overcome the transverse optical force pulling the cell into the optical gradient, the escape velocity, can be measured for each cell and accumulated for a given population to provide a signature for that population. To achieve a higher throughput of cell analysis, the fast-scan technique, which allows analysis of multiple cells simultaneously, was developed.[1]

7.2.1 The Fast-Scan Technique

To maximize measurement throughput, we have developed a technique in which many cells can be analyzed in parallel as shown in Figure 7.1. Particles or cells are prepared in suspension and then dispersed on the surface of an analysis slide. A long, thin laser line is then generated by a near-infrared (IR) laser and used to scan quickly across the optical field of view. Cells in proximity of the laser will experience an attractive force due to the high cross-sectional gradient of intensity of the line. As the laser line passes over each cell, the optical forces first cause the cell to move backward toward the oncoming line, then be pulled forward as the line continues moving forward. Because the cell spends more time interacting with the beam as it moves forward, there is a net displacement of the cell in the forward direction. The scanning speed of the laser is chosen to be sufficiently high so that cells are only incrementally displaced and are not caught up by the laser and swept out of the field of view. The cell is usually displaced only a few microns by the passing of the laser, therefore multiple scans of the laser (typically 5 to 7 scans) are used to increase the total displacement and enhance the detection sensitivity. The net displacement of each cell is measured by tracking the positions of the cells after every scan, using custom-image-processing software.

The measured displacement value of the cell correlates to both the optical force exerted on the cell as well as the fluidic drag and surface interaction forces that oppose the cell's

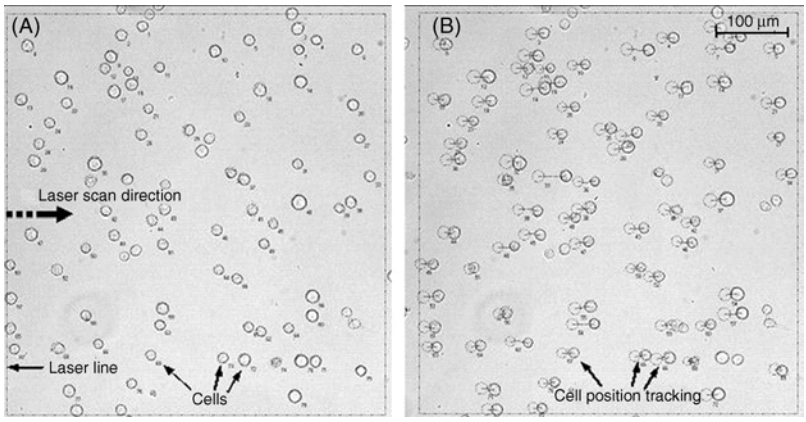


FIGURE 7.1
The fast-scan technique optimizes measurement throughput by analyzing many cells in parallel.

movement. Changes in the overall size, morphology, or refractive index of the cell will alter the magnitude of the optically induced forces, whereas changes in the size, cell membrane, or structural rigidity will affect the drag and surface interaction forces. All of these forces are a function of the composition and structure of the cell and, therefore, correspond to a measurement of the biological state of the cell. In addition, by having a laser line that scans across the cell, this technique emphasizes an integrated measurement of the whole cell rather than focusing on a single organelle or the membrane as is usually the case in experiments performed with laser tweezers. As a result, this measurement technique can be general in use and applied to many different cell types for multiple applications.

7.2.2 The Fast-Scan Instrument

A schematic of the laser scanning instrument is shown in Figure 7.2. A continuous-wave (CW) ytterbium fiber laser was chosen as the laser source because of the high-CW powers available (this laser has a maximum output of 20 W, linearly polarized) and the diffraction-limited output beam from the single-mode fiber delivery. The wavelength of the laser is 1070 nm, which is below the high-IR absorption band of water but above the ultraviolet–visible absorption band of cells. The laser line (760-µm long at the sample plane) is generated along one axis using a flexure-type galvanometer oscillating at 100 Hz. Two relay lenses image the deflected beam from the galvanometer to the entrance pupil of the microscope objective, which is a 20×, 0.4 numerical aperture (NA), near-IR, infinity-corrected lens. A variable beam expander determines the diameter of the collimated beam entering the microscope objective, which in this case was set to 4.1 mm. The effective NA of the focused laser is, therefore, smaller than that of the microscope objective because the entrance pupil of the objective is under-filled. It was observed experimentally that although having a more highly focused beam in general increases the lateral optical force imparted on the particle, this increase starts to flatten out above an NA of 0.2. Because the tolerances on focusing and alignment of the optical system can be relaxed by having a lower NA, we chose to use NA = 0.2 for these experiments. The resulting spot size for the focused beam is ~6 µm (1×10^{-2} diameter). We have chosen to have the laser illuminate the sample from the bottom to allow a clear, unobstructed view of the cells, although we have previously observed that the laser can

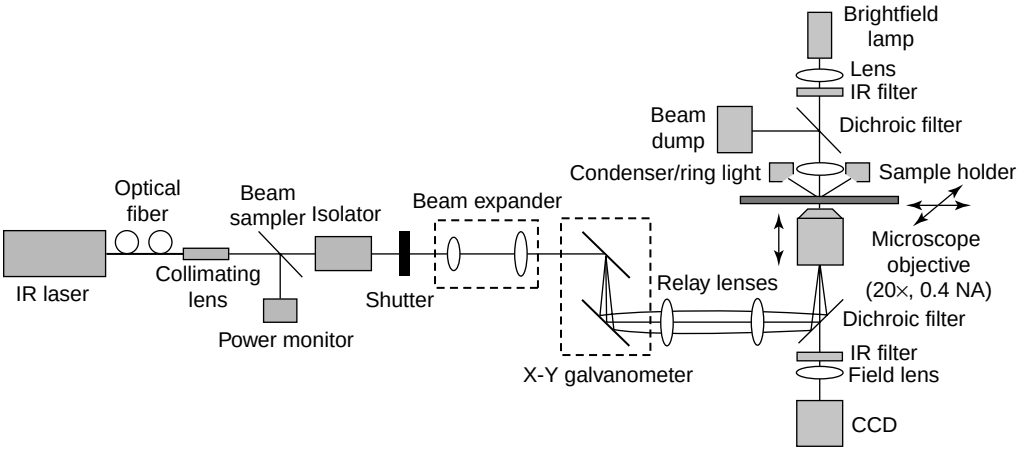


FIGURE 7.2
A schematic of the laser scanning instrument.

also illuminate the sample from the top with similar performance. The laser is focused by moving the microscope objective up and down axially on a stepper motor stage. The sample holder is mounted on two orthogonal piezoceramic-driven linear stages that allow precision movements with velocity regulation better than ±5% measured at a 10-Hz sampling rate. A dichroic mirror below the microscope objective allows imaging of the sample to a 10-bit, megapixel charge coupled device (CCD). An f = 150-mm field lens (PAC086, Newport Corporation, Irvine, CA, U.S.A.) provides an imaging magnification of 15× resulting in a 600 × 600-μm field of view at the sample plane being observed on the CCD. Either bright- or dark-field illumination is permitted by white light sources mounted above the sample.

7.2.3 Characterization of the Instrument

Although the optical forces and fluid drag forces induced on the particle are mostly linear with incident laser power and scanning velocity, respectively, the total displacement of the particle will be nonlinear as a function of these variables because of the dynamics of the particle as the laser passes over it. Figure 7.3 shows the displacement of 15-μm polystyrene beads as a function of both of these variables: in one case, keeping a constant scanning velocity and, in the other, a constant power. As the power increases or the velocity decreases the particle displacement becomes larger and eventually approaches a threshold after which the particle becomes caught and continues moving with the laser line. The measured displacement of a particle will remain constant if the power of the laser and scanning velocity are scaled proportionately (i.e., power/velocity = constant).

Measurement of a homogeneous population of calibrated beads provides an indication of the instrument noise (see Figure 7.4). For a laser power of 2.5 W, a scanning velocity of 110 μm/sec, five scans of the laser, and a total of 324 particles measured, we find a mean displacement of 16.9 μm and a coefficient of variation (CV) of 13.3%. Some of the most significant noise sources that contribute to this variation include laser power fluctuations, stage velocity noise, accuracy of the particle position determined by the image processing, inhomogeneities of the particles, surface nonuniformities, temperature variations, and Brownian motion of the particles. The asymmetrical profile of the distribution results from

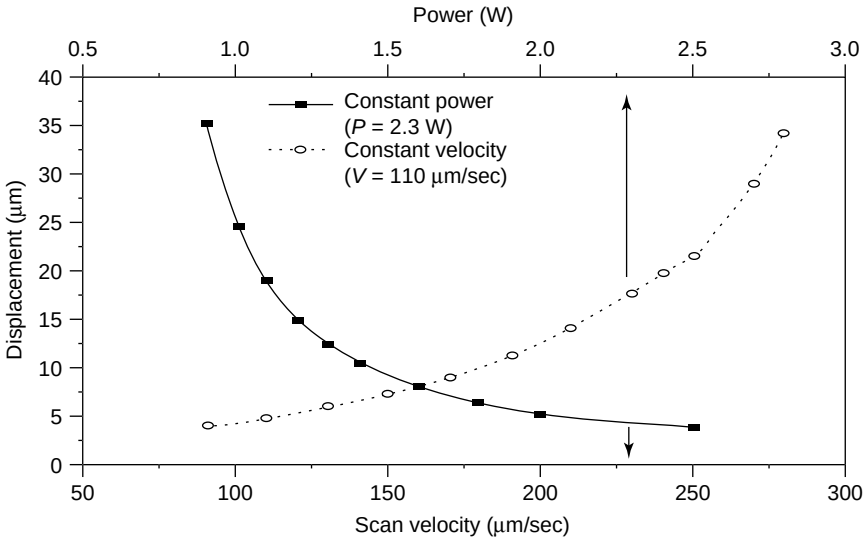


FIGURE 7.3
Displacement of 15-µm polystyrene beads as a function of either constant scanning velocity or constant laser power.

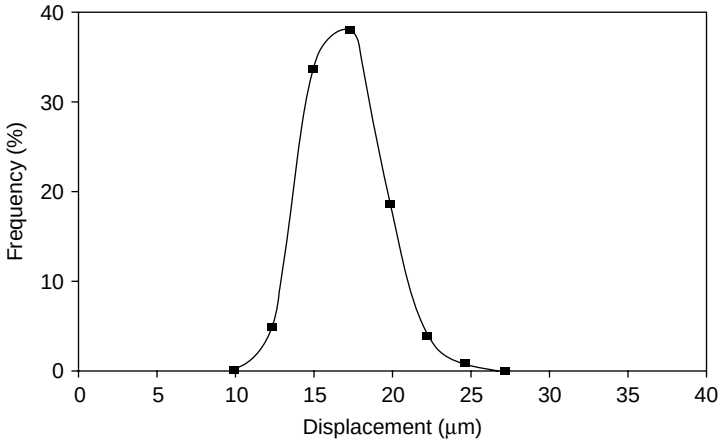


FIGURE 7.4
Measurement of a homogeneous population of calibrated beads provides an indication of the instrument noise.

the nonlinear response of the particles to the laser. Nonnormal distributions are even more pronounced for populations with larger inherent variations.

The sensitivity of our measurement technique to particle size was evaluated using a size standard set of 5-, 7-, 10-, 15-, and 20-µm-diameter beads. The mean displacement of each population is shown in Figure 7.5. In these measurements, the 5- and 7-µm-diameter beads were too light and were thus pushed vertically off the surface of the slide resulting in longer displacements because of reduced surface friction and fluidic drag forces. These measurements were also noisier because of the intermittent contact of the beads with the surface. For cell measurements, we prefer that the cells do not leave the surface during measurement. Fortunately, the vertical scattering force is less on cells, and we are routinely able to determine measurement parameters to allow analysis of cells as small as 5 µm or

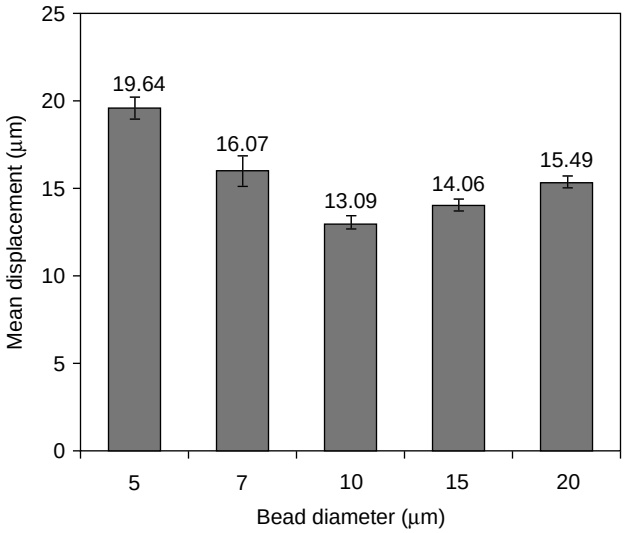


FIGURE 7.5
Measurement sensitivity to particle size using 5-, 7-, 10-, 15-, and 20-μm-diameter beads; mean displacement shows 5- and 7-μm-diameter beads were too light and also noisier; for diameters between 10 and 20 μm, the beads show a slowly increasing trend in displacement; a relatively low size-sensitivity results for particles in this size range is advantageous for measuring cells that often have widely varying sizes even in otherwise homogenous populations.

smaller. For diameters between 10 and 20 μm, the beads stayed on the surface during measurement, and we see a slowly increasing trend in displacement. This increase is a result of the increasing optical interaction mass for larger particles; however, the rate of increase is low because this increasing optical force is being counteracted by an also-increasing drag force due to the larger diameter of the particle. A relatively low size sensitivity results for particles in this size range, which is advantageous for measuring cells that often have widely varying sizes even in otherwise homogenous populations. We have observed experimentally with cells that usually only a shallow dependence on cell size results for cells within a single population.

In general, measuring beads requires less laser power than measuring cells because of their higher index of refraction ($n \sim 1.5$ for polystyrene beads vs. $n \sim 1.37$ for cells).[15] The optical force imparted to a particle scales with the difference in index of refraction between the particle and the fluidic medium.[16] For bead measurements, we typically operate at a laser power of 2.5 W, whereas for cell measurements the laser is operated at 10 W to obtain similar displacements. These relative power levels are in line with the comparative refractive index differences between the two different particle types and water.

7.3 Biological Examples of the Uses of Optophoresis

To demonstrate the utility of optophoresis in the preclinical stages of drug discovery, we screened several cell lines for their response to three standard chemotherapy drugs: fludarabine, vincristine, and Gleevec (imatinib mesylate, Novartis Pharmaceuticals, Basel, Switzerland) (Figure 7.6). To perform these measurements with the fast-scan technique, cells were suspended in a 9-mm-diameter well in a sample holder. The bottom surface of

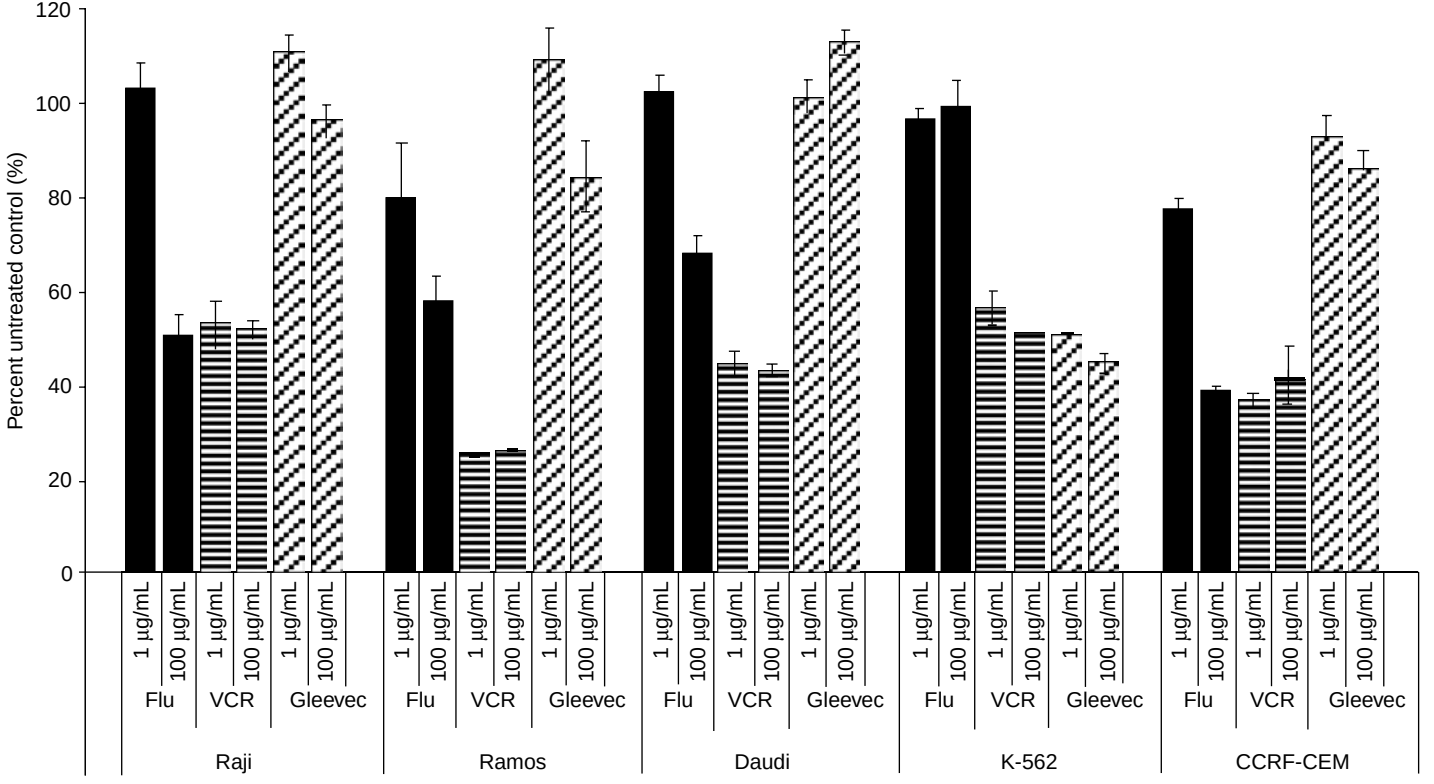


FIGURE 7.6
Several cell lines were screened for their response to three standard chemotherapy drugs: fludarabine, vincristine, and Gleevec; decreases in the mean optophoretic distance of Bcr-Abl positive cell line K-562 with no effect in the other cell lines demonstrate that the downshift in mean displacement is specific to the presence of the Bcr-Abl oncoprotein; fludarabine, however, was ineffective in causing a decrease in the optophoretic displacement of the K-562 cell line, but showed an effect in the other four cell lines; mobility values for each cell in the population have been taken and averaged, but data on subsets of the cells can still be extracted and further analyzed.

the well was a glass coverslip coated with agarose (Sigma-Aldrich, St. Louis, MO, U.S.A.) by spin deposition (~0.1 μm dry thickness) (Headway Research, Inc., Garland, TX, U.S.A.) to prevent adhesion of the cells. A 60-μm spacer separated the top glass coverslip from the bottom surface, enabling the use of small sample volumes of about 4 μL per sample. A line-shaped optical gradient was produced by focusing a raster-scanned laser beam into the well. A translation stage moved the well and, consequently, all the cells in the field of view, past this stationary laser line at a constant velocity. The displacements of all the cells in the field of view were monitored by a CCD camera. The translation velocity determined how far the cells were displaced. Slower translation velocities displaced the cells to a greater extent than faster velocities. Cells were typically exposed to the moving gradient multiple times to increase the total distance traveled and improve the signal resolution. The position and displacement of each cell in the field of view was determined by software image processing. From these data, the distance that each cell travels was measured. Multiple fields of view were measured to accumulate an appropriate number of cells for each population, typically 400 cells. The distribution of displacements for each cell population was generated, and the mean optophoretic distance was calculated to characterize that population. Different populations were then compared in terms of their mean optophoretic distances measured under the same experimental conditions.

All five cell lines responded to vincristine, a nonspecific agent that interferes with the cytoskeleton (Figure 7.6). On the other hand, decreases in the mean optophoretic distance of Bcr-Abl positive cell line K-562 with no effect in the other cell lines in the presence of 1 μM Gleevec after 48 h of treatment demonstrate that the drug-dependent optophoretic downshift in mean displacement is specific to the presence of the Bcr-Abl oncoprotein. Conversely, fludarabine, a nucleoside analog was ineffective in causing a decrease in the optophoretic displacement of the K-562 cell line, but showed an effect in the other four cell lines. Activity of Bcr-Abl in the K-562 cell line has been previously shown to overcome the inhibitory effect of fludarabine.[17]

In addition, the drug responses of these same cell lines were measured with a more traditional measurement technique, the Vialight assay (Cambrex, East Rutherford, NJ, U.S.A.), which quantifies the intracellular levels of adenosine triphosphate (ATP) (Figure 7.7). A complete agreement can be seen between the data generated using optophoresis and that by the Vialight method; however, several differences in these two techniques are worth noting. The Vialight method is destructive, requiring lysis of the cells and giving an average value for the entire population. Optophoresis, on the other hand, is nondestructive to the cells. In the experiment shown in Figure 7.6, the mobility values for each cell in the population have been taken and averaged, but data on subsets of the cells can be extracted and further analyzed as well.

Dose–response characterization is a critical step in the optimization and comparison of the potency of different compounds. In Figure 7.8, two human cell lines, the Daudi and Raji lines, with very different sensitivities to the drug fludarabine were measured with optophoresis over a 3 log range of concentrations. The mean displacement value for each drug concentration was plotted and yielded a sigmoid dose–response curve. The Raji cell line had an EC_{50} value of 1.1 ± 0.27 μg/mL, whereas the much less-sensitive Daudi line had an EC_{50} of 11.51 ± 3.2 μg/mL using optophoresis. Almost identical EC_{50}s were obtained with the Vialight method, that is, they were 1.62 μg/mL for Raji and 11.27 μg/mL for Daudi.

Although cell lines are often used in preclinical studies, it is also important to test a compound on primary cells from the target tissue itself. Because this tissue must be taken from healthy, human volunteers, such tissue is often difficult to obtain in sufficient quantities for many types of assays that require large numbers of cells. In Figure 7.9, normal peripheral blood mononuclear cells (PBMCs) from four unrelated, healthy donors were treated

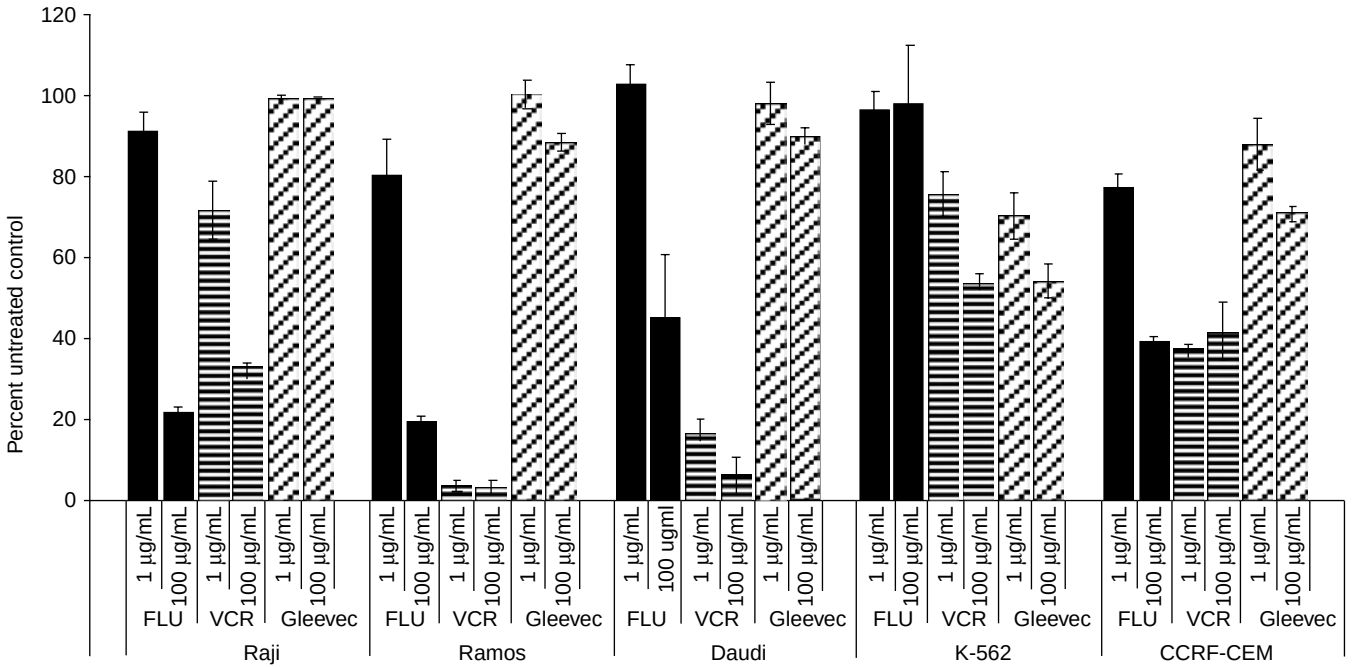


FIGURE 7.7
Drug responses of the same cell lines using optophoresis and Vialight were compared; complete agreement can be seen between the data generated by both methods.

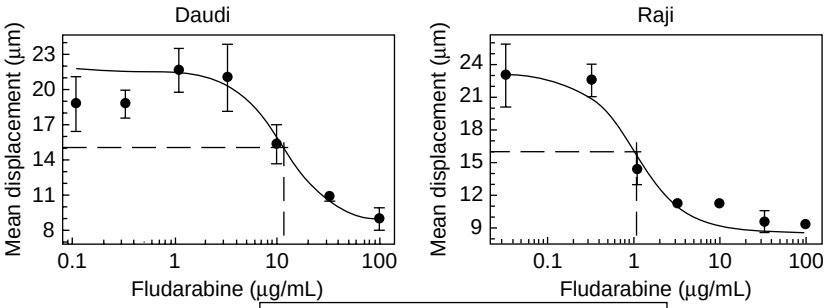


FIGURE 7.8
The Daudi and Raji human cell lines, with different sensitivities to fludarabine were measured with optophoresis over a 3 log range of concentrations; the Raji cell line had an EC_{50} value of 1.1±0.27 μg/mL, whereas the much less-sensitive Daudi line had an EC_{50} of 11.51±3.2 μg/mL.

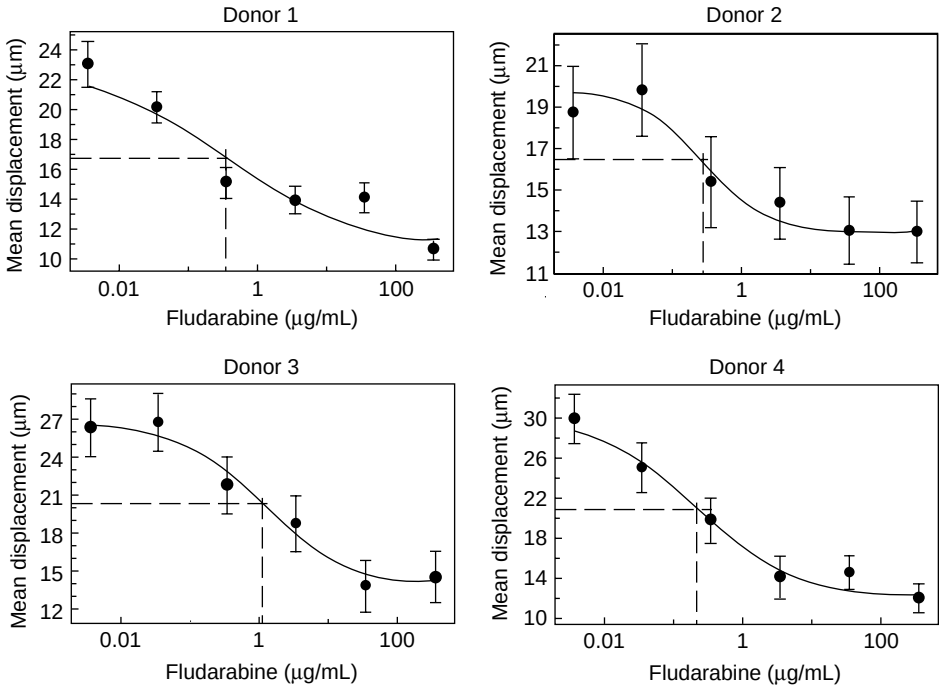
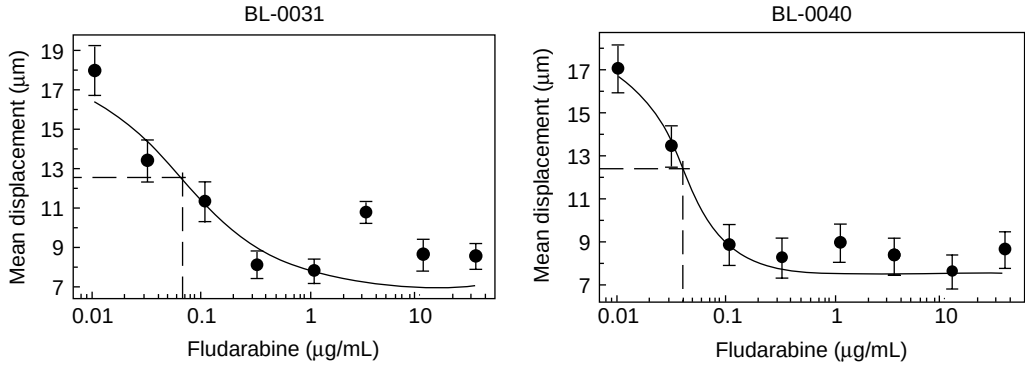


FIGURE 7.9
Normal peripheral blood mononuclear cells (PBMCs) from four unrelated, healthy donors treated with fludarabine; all four of the control individuals had almost identical dose–response curves.

in vitro for 48 h with fludarabine. The displacement values for the donor's PBMC are shown over a range of 2.5 log concentrations of the drug. All four of the control individuals had almost identical dose–response curves to fludarabine as measured using optophoresis, with EC_{50} values around 1 μg/mL.



	EC_{50} (μg/mL)	
	Optophoresis	Vialight
BL-0031	0.07 ± 0.03	0.08 ± 0.014
BL-0040	0.04 ± 0.007	0.07 ± 0.015

FIGURE 7.10
The optophoretic response of the cells from two chronic lymphocytic leukemia (CLL) patients, showing fludarabine EC_{50} values below 0.1 μg/mL, which is over a log lower than the EC_{50} values obtained for the four control individuals shown in Figure 7.9, demonstrating the increased sensitivity of the CLL tumor cells to fludarabine relative to normal cells exposed to the drug.

Patients need to be monitored for a drug's effects both during clinical trials and often after the drug has been approved, in measurement of the drugs effectiveness. The Optophoresis assay can be performed very easily on tissues obtained from clinical trial participants or from patients. Because a very small number of cells is needed, even a biopsy may be sufficient to give enough cells for a full dose–response curve.

Continuing with the example of fludarabine, in patients with chronic lymphocytic leukemia (CLL), we have measured the optophoretic shift from a blood sample. In these samples, over 90% of the cells were tumor cells, as evidenced by fluorescence-activated cell sorter (FACS) analysis of CD20/CD19 expression; therefore, the EC_{50} that is obtained is a reflection of the response of the tumor cells and not the normal cells of the blood. Figure 7.10 shows the optophoretic response of the cells from two CLL patients to fludarabine over a 3 log range of concentrations of the drug. Both patients had fludarabine EC_{50} values below 0.1 μg/mL, which is over a log lower than the EC_{50} values obtained for the four control individuals shown in Figure 7.9, demonstrating the increased sensitivity of the CLL tumor cells to fludarabine relative to normal cells exposed to the drug.

7.4 Discussion

Optophoresis is a technology that provides sensitive, quantitative measurements of cellular responses, requires little sample processing, and uses no external labels or additional reagents. In addition, optophoresis requires very few cells to yield statistically significant and robust results. For example, the experimental error bars for all the data presented are comparable between the metabolic activity assay, which requires approximately 60,000 cells and optophoresis, which requires about 2,000 cells per well. This

reduced sample requirement has a significant impact on difficult-to-obtain cells, enabling, among other things, assays from patient materials in clinical settings. Although the current study demonstrates the utility of optophoresis to quantitate cellular responses to fludarabine, we have applied this technology to analyze both suspension and adherent cell lines, as well as solid tissues for evaluation of cellular responses other than apoptosis.

Optophoresis was developed to use the forces generated by moving optical gradients to quantitatively measure cellular responses. This technique is distinguished from all of the previous optical trapping techniques in that the cells are not trapped but, instead, have a quantifiable, characteristic motion induced by the passage of an optical gradient over the cells. To our knowledge, no other technique performs such a quantitative, whole-cell measurement on live cells without additional cell manipulation.

Sensitivity of patients' cells to a chemotherapeutic drug has been measured in the past using several types of technologies with mixed results. Clonal growth assays have the limitation that whereas resistance is easy to measure by the presence of growth, levels of sensitivity are less quantifiable. In addition, the very process of growing cells out from a sampling of tumor cells can lead to the artifact of a small number of cells from a subpopulation outgrowing the rest of the cells over a several week period and giving a skewed result to the overall interpretation of the data.

On account of the features of optophoresis, it can be used throughout the various stages of drug discovery, in clinical trials, and even in the diagnostic arena for predicting which therapy may be optimal for a given patient. It may also be possible to measure *ex vivo* the cytotoxic effect of the drug on the normal cell population in a patient sample. By comparing that value with the EC_{50} of the tumor cell population, one can determine which drug has the largest safety index between the desired killing of tumor cells and the undesired killing of the normal cells of the patient. We are continuing to expand the uses and roles of optophoresis in drug development, drug discovery, and clinical diagnostics.

References

1. Wang, M.M., Schnabel, C.A., Chachisvilis, M., Yang, R., Paliotti, M.J., Simons, L.A., McMullin, L., Hagen, N., Lykstad, K., Tu, E., Pestana, L.M., Sur, S., Zhang, H., Butler, W.F., Kariv, I., and Marchand, P.J., Optical forces for noninvasive cellular analysis, *Appl. Opt.,* 42, 2003, 5765–5773.
2. Forster, A.H., Wang, M.M., Butler, W.F., Chachisvilis, M., Chung, T.D., Esener, S.C., Hall, J.M., Kibar, O., Lykstad, K., Marchand, P.J., Mercer, E.M., Pestana, L.M., Sur, S., Tu, E., Yang, R., Zhang, H., and Kariv, I., Use of moving optical gradient fields for analysis of apoptotic cellular responses in a chronic myeloid leukemia cell model, *Anal. Biochem.,* 327, 2004, 14–22.
3. Ashkin, A., Acceleration and trapping of particles by radiation pressure, *Phys. Rev. Lett.,* 24, 1970, 156–159.
4. Ashkin, A., Dziedzic, J.M., Bjorkholm, J.E., and Steven, C., Observation of a single-beam gradient force optical trap for dielectric particles, *Opt. Lett.,* 11, 1986, 288–290.
5. Arai, Y., Yasuda, R., Akashi, K., Harada, Y., Miyata, H., Kinosita, K., Jr., and Itoh, H., Tying a molecular knot with optical tweezers, *Nature,* 399, 1999, 446–448.
6. Ashkin, A., Dziedzic, J.M., and Yamane, T., Optical trapping and manipulation of single cells using infrared laser beams, *Nature,* 330, 1987, 769–771.
7. Ashkin, A. and Dziedzic, J.M., Optical trapping and manipulation of viruses and bacteria, *Science,* 235, 1987, 1517–1520.

8. Svoboda, K. and Block, S.M., Biological applications of optical forces, in *Annual Review of Biophysical and Biomolecular Structure,* Vol. 23, Stroud, R.M., Cantor, C.R., and Pollard, T.D., Eds., Annual Reviews Inc., Palo Alto, CA, 1994, pp. 247–285.
9. Svoboda, K., Schmidt, C.F., Schnapp, B.J., and Block, S.M., Direct observation of kinesin stepping by optical trapping interferometry, *Nature,* 365, 1993, 721–727.
10. Wang, M.D., Yin, H., Landick, R., Gelles, J., and Block, S.M., Stretching DNA with optical tweezers, *Biophys. J.,* 72, 1997, 1335–1346.
11. Tadir, Y., Wright, W.H., Vafa, O., Ord, T., Asch, R.H., and Berns, M.W., Force generated by human sperm correlated to velocity and determined using a laser generated optical trap, *Fertil. Steril.,* 53, 1990, 944–947.
12. Patrizio, P., Liu, Y., Sonek, G.J., Berns, M.W., and Tadir, Y., Effect of pentoxifylliine on the intrinsic swimming forces of human sperm assessed by optical tweezers, *J. Androl.,* 21, 2000, 753–756.
13. Svoboda, K., Schmidt, C.F., Branton, D., and Block, S.M., Conformation and elasticity of the isolated red blood cell membrane skeleton, *Biophys. J.,* 63, 1992, 784–793.
14. Sleep, J., Wilson, D., Simmons, R., and Gratzer, W., Elasticity of the red cell membrane and its relation to hemolytic disorders: an optical tweezers study, *Biophys. J.,* 63, 1999, 3085–3095.
15. Drezek, R., Dunn, A., and Richards-Kortum, R., Light scattering from cells: finite-difference time-domain simulations and goniometric measurements, *Appl. Opt.,* 38, 1999, 3651–3661.
16. Ashkin, A., Forces of a single-beam gradient laser trap on a dielectric sphere in the ray optics regime, *Biophys. J.,* 61, 1992, 569–582.
17. Robertson, L.E., Denny, A.W., Huh, Y.O., Plunkett, W., Keating, M.J., and Nelson, J.A., Natural killer cell activity in chronic lymphocytic leukemia patients treated with fludarabine, *Cancer Chemother. Pharmacol.,* 37, 1996, 445–450.

8

Patient-Derived Primary Cells in High-Throughput Differential Antitumor Screens: Let the Patients Be the Guide

Irwin A. Braude

CONTENTS

8.1 Introduction

Paul Ehrlich's pioneering work in 1905, which culminated in the discovery of arsephenamine (also known as compound 606 or Salvarsan) for the treatment of syphilis,[1] launched the modern concept of using drug discovery screens to identify chemotherapeutics. This approach of screening hundreds, and later thousands, of compounds in a "bioassay" was subsequently applied to the discovery of antitumor drugs. Between 1947 and 1955, Cornelius Packard "Dusty" Rhoads and his team at Sloan-Kettering Institute (New York, NY, U.S.A.) screened an estimated 20,000 compounds (both synthetic and natural) in a sarcoma-180 mouse tumor model.[2] Around the same period, Murray Shear and his colleagues at the U.S. National Cancer Institute (NCI, Bethesda, MD, U.S.A.) tested 3000 chemicals and plant materials for potential antitumor activity.[3] Later, Abbott, Hartwell, and others at the NCI screened a wide range of natural products in both *in vitro* and *in vivo* antitumor bioassays.[4]

As early as the 1950s, the value of employing primary cells derived from patients' tumors in an antitumor drug discovery screen was recognized.[5] The use of patient-derived diseased tissue proved impractical at that time, however, because of the inability to maintain some of the cells in culture and because of insufficient tumor tissue available from patients undergoing surgery. In the 1980s, researchers at the NCI attempted to use patient-derived diseased tissue in a high-throughput screening (HTS) format using a variation of the Hamburger and Salmon clonogenic chemosensitivity assay.[6] The authors reported that

the approach yielded some useful results but was deemed impractical in an HTS.[7] Scientists at pharmaceutical companies met similar disappointment while attempting to use patient-derived primary cells in an HTS. The trend then shifted to bioassays using "cell lines" because of their ease of use,[8] and later, "mechanistic" screens based on "validated" biochemical targets became and remained the HTS format of choice.

Cell line and mechanistic screens offer many practical advantages for HTS, but the trade-off is a lack of clinical relevance. Although bioassays using cell lines have the ability to use the whole cell to detect the effect of a compound on any of the thousands of biochemical events occurring within it these cells have become laboratory adaptations of the original tumor cell population and lack the heterogeneity found in primary tumor cells. Mechanistic screens, which are inherently limited to screening for one target at a time, face the problematic challenge of identifying and validating clinically relevant targets because, even though identified targets may be associated with the disease under investigation, antagonizing or agonizing the targets may not benefit the patient.

This chapter describes a screen that uses primary diseased and normal cells in a high-throughput differential format that offers the advantages of both whole-cell bioassays and mechanistic screens. The Cytection technology can be a valuable tool in identifying new and more effective drugs to treat cancers and other diseases. This two-part screen consists of a cell processing and isolation technology and a rigorous, stringent four-level screen that scrutinizes compounds for activity, potency, selectivity, and spectrum of activity. The screen's applicability includes drug discovery screens where well-validated targets are not apparent, patient tissues are legally and readily available, and bioassay endpoints reflect clinically desirable goals.

This approach offers two opportunities to discover clinically relevant compounds. The first is the compounds identified directly in the Cytection screens. Second, appreciating that these compounds have the desirable biological endpoints, they can be used as "molecular probes" to determine their putative mechanism(s) of action. This "reverse drug discovery"[9] identifies "validated" targets that can be the basis for mechanistic screens that could lead to the discovery of additional compounds.

Although this approach and criteria are applicable to many diseases (in particular autoimmune and infectious diseases), its use in antitumor screens is described below.

8.2 The Processing and Use of Patient-Derived Primary Cells

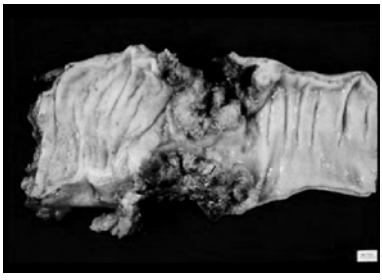
Both tumor and adjacent normal tissues are provided by a nationwide network of hospitals after approval by their respective Institutional Review Boards. The specimens undergo pathological examination at the source and are minced and shipped via overnight express on wet ice. Upon arrival, tissues are finely minced with opposing scalpels and then subjected to one or more rounds of digestion employing a cocktail of enzymes. Microscopic examination determines when sufficient clusters of cells are liberated from the intact tissue. The liberated cells are separated from the residual intact tissue and fresh enzyme solution is added to the remaining undigested tissue. The process is repeated until no further cell clusters are isolated.

Tumor cells are separated from normal cells by taking advantage of their differences in size, density, and antigenicity. As there are few, if any, well-characterized tumor-specific antigens, negative immunoselection, employing monoclonal antibodies directed against normal cell-surface antigens, is employed. The harvested cells are then aliquoted into

ampules, frozen in a programed cell freezer, and stored in the vapor phase of liquid nitrogen cell-storage freezers.

As a quality control measure, several segments of the predigested tissue are fixed and sent to a histopathologist to confirm tumor diagnosis and grading. Where there is agreement between the external pathologist's conclusion and the findings of the pathologist at the hospital that offered the specimen, the information is archived for future reference. Where there is disagreement, at least two additional histopathologists are consulted, and the consensus determination is archived. The pathologists also examine the adjacent normal tissue to confirm that it is free of any detectable tumor cells.

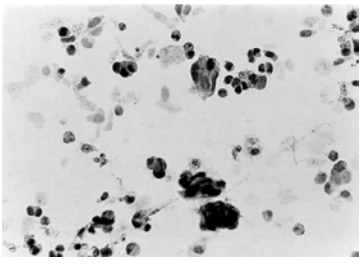
As a further quality control measure, after the tissue digestion and cell isolation procedure is complete, approximately 20,000 to 30,000 clusters of cells are cytospun, fixed, and sent to cytopathologists for examination. Cytopathologists determine and document the percentage of tumor cells and normal cells present in a preparation and the composition of the normal cell population. Figure 8.1 shows a picture of a colon segment harboring a carcinoma tumor. In the same figure, the micrographs illustrate the



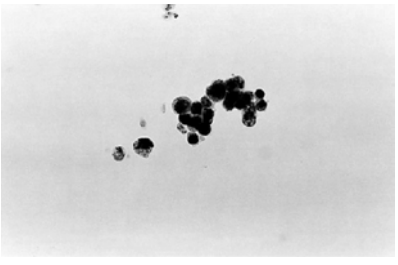
Colon

Data recorded:

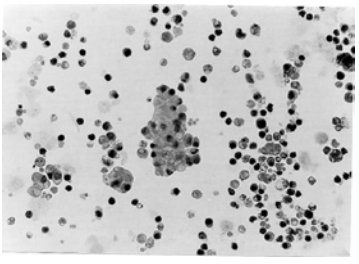
- Patient histories
- Tissue source
- Grading of tumor
- Normal tissue type
- Sample storage data



Colon tumor cluster



Colon metastases



Normal colon cells

FIGURE 8.1
Cytection generates heterogeneous cell populations from patient tissue for screening.

morphological differences between populations of tumor cells and normal cells: the heterogeneity of the tumor cell population in the primary tumor sample, the lesser amount of heterogeneity in the metastatic (liver) tumor, and the normal cell population composed of both crypts and inflammatory cells. Also noteworthy is the formation of clusters or "tumoroids," which is typical of carcinomas.

This digestion, isolation, and storage approach has been successfully applied to a wide range of tumor and normal tissues. Tumor tissues include both primary and metastatic tumors of the more common types, such as colon, lung, and ovarian tumors, melanomas, and sarcomas, and rare tumors, such as schwannoma. Normal cell populations include lung, ovary, colon, heart, liver, kidney, and blood.

The bioassay portion of Cytection consists of thawing the cells and removing the cryopreservative by centrifugation. A cell suspension is made and is added to wells of specially designed submicrotiter trays, precoated with a human extracellular matrix (ECM). The submicrotiter plates hold a maximum volume of about 50 μL and, unlike commercially available versions, are designed so that no evaporation of the media occurs for 21 days, and there are no "edge effects" along the periphery of the plate. The ECM is similar to the substrate found *in vivo* and allows for good cell plating and seeding efficiency. The tissue culture media is serum-free, and its composition is tailored for each cell type.

The cells are permitted to "plant" to the ECM and adjust to the incubator temperature (37°C) and CO_2 concentration. Then test compounds or controls (both in 0.1% dimethyl sulfoxide, DMSO) are added to the test wells. The cells are then incubated overnight, and the indicator dye Alamar blue[10] is added. This noncytotoxic dye reacts to mitochondrial redox reactions and is measured fluorometrically. Cell metabolic activity is determined starting at 3 h after the dye is added and daily thereafter.

Before using patients' cells in a screen, the bioassay is calibrated to determine the number of cell clusters (see above) required to achieve the desired signal above the media controls not containing cells (i.e., the signal-to-noise ratio). Figure 8.2 shows an example of a typical calibration. Eighteen million clusters of tumor cells were harvested from this patient's lung adenocarcinoma. The cells were seeded at concentrations ranging from 330 to 10,000 cell clusters per test well. The metabolic activity was then determined at different time intervals for up to 10 days. It was observed that the majority of the dye conversion occurred within the first 3 to 5 days of incubation. It was also determined that 1000 clusters/well yielded the desired minimum threefold signal-to-noise ratio. Therefore, it was calculated that 18,000 test wells could be used from this single patient's tumor. The historical average yield for carcinoma tumors and soft tissue sarcomas processed was approximately 20,000 and 170,000 test wells per tumor tissue, respectively.

For purposes of quality control, any cell preparation used in the screen (see below) needs to have a definitive or consensus diagnosis and grading determined by histological examination. Furthermore, tumor cell preparations must be at least 80% tumor cells (as determined by cytopathological examination), whereas normal cell preparations must be devoid of any tumor cells. Cell viability, determined by trypan-blue exclusion, must be a minimum of 70%, and the signal-to-noise ratios for a predetermined cell cluster concentration must be at least threefold.

Based on the results from those cell calibration assays, the routine incubation period was set at 3 days. During that interval, some of the cells will proliferate, some will be quiescent but metabolically active, and some will die, which is the case for most primary cell cultures. Because these cells are cultured for periods far shorter than their typical population doubling time (about 6 to 7 days), these cells should not fully adapt to the tissue culture conditions and, therefore, maintain their primary cell phenotype.

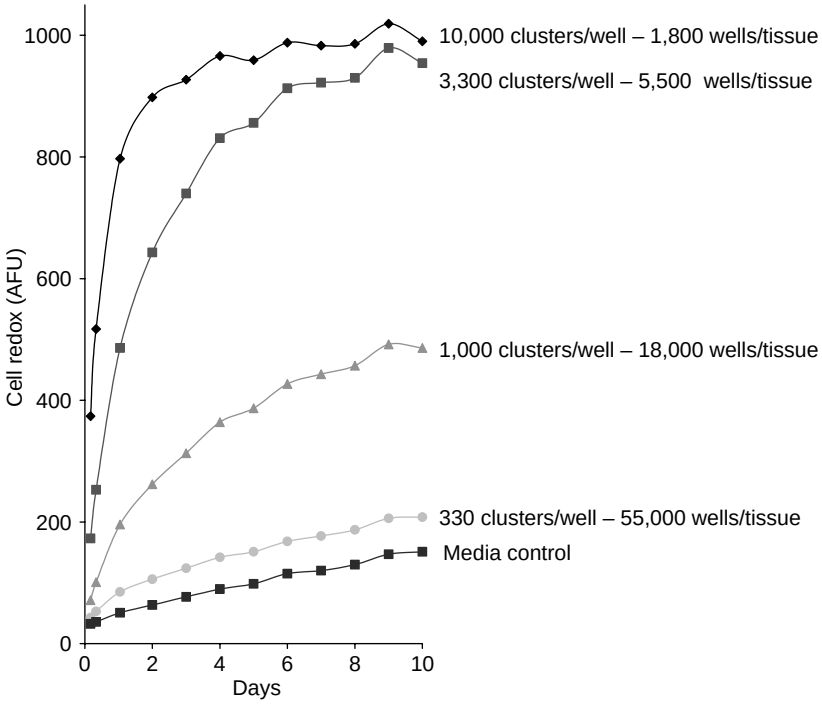


FIGURE 8.2
Antitumor assay calibration for lung adenocarcinoma.

8.3 The Screen

The antitumor screen involves four levels of testing, each of greater rigor and stringency. The goal is to identify which compounds are the most potent, the most selective for tumor vs. normal cells and which show the greatest spectrum of activity against tumor cells but not against normal cells.

Before compound collections are screened, the collection is calibrated to determine the concentration at which it should be tested. Five hundred samples that are representative of the entire collection (single compounds, pools of combinatorial compounds, or extracts) are serially diluted to identify the concentration at which 2 to 3% of the compounds show at least 80% inhibition of tumor cells within two to three standard deviations from the mean. This concentration is then used to screen the whole collection.

As illustrated in Figure 8.3, samples are tested on cells from one patient's tumor in Level 1 of the screen to determine globally which samples exhibit the greatest activity at the predetermined concentration. Samples are tested without replicates. Typically, the throughput for the screen is 4000 to 6000 test wells per run set; it can, however, be easily scaled to 10,000 or more test wells per run set.

Samples that are advanced to the next level of the screen are selected on the same basis as the calibration: those inhibiting at least 80% of the tumor cells within two to three standard deviations from the mean for a one-run set. Figure 8.4 shows a typical "scatter plot" for 3680 compounds tested in a run set. The mean inhibition for all the compounds screened was 7.7%. The horizontal bars represent the standard deviations above and below that mean. The 105 compounds deemed "active" (shown by the bold bar in the

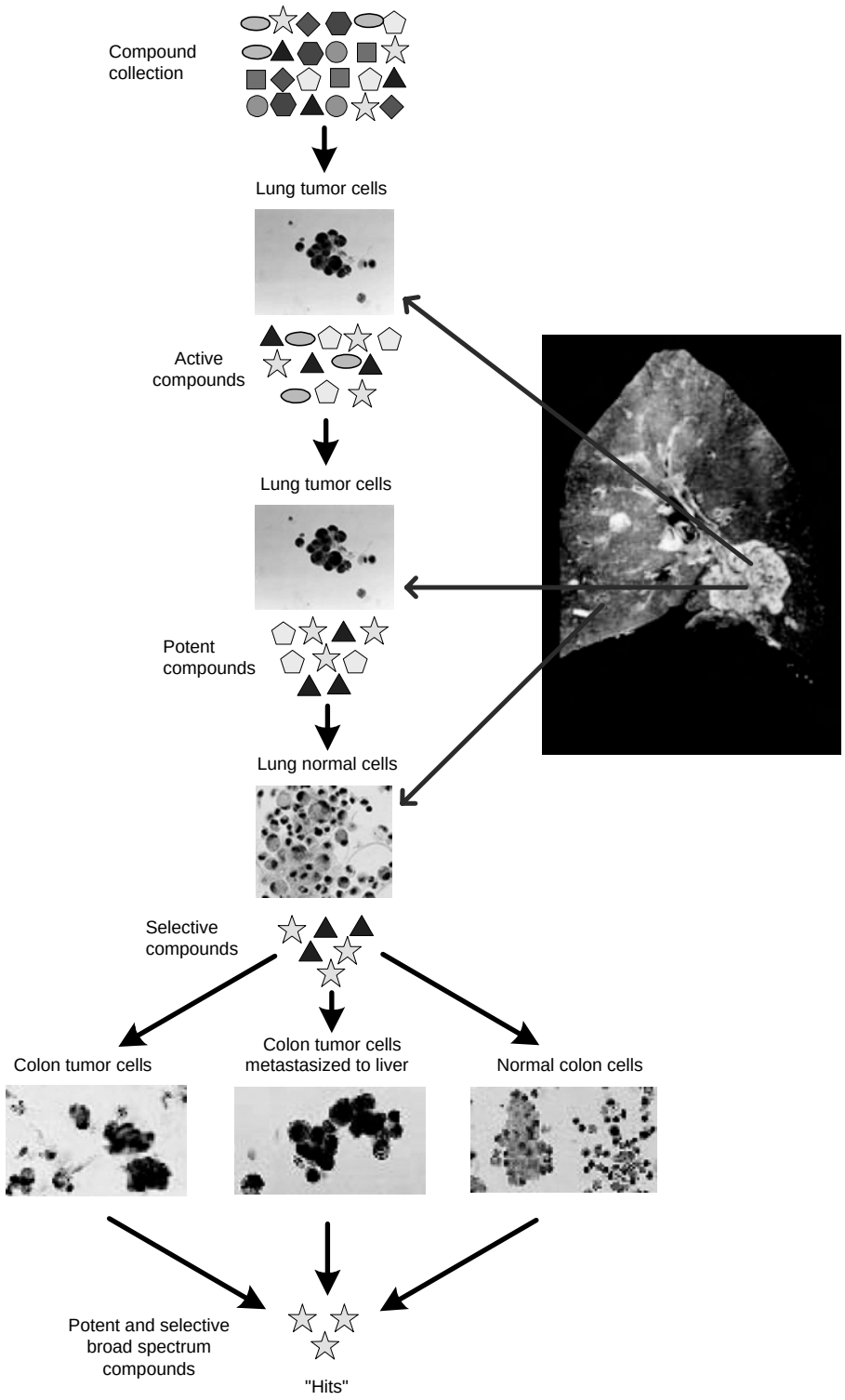


FIGURE 8.3
Compounds undergo four levels of screening on patient-derived primary cells.

graph) have an average inhibition of 87% and are greater than two standard deviations above the mean. They represent 2.9% of the compounds tested. Compounds that have the desirable bioactivity but that are judged by chemists as potentially being overtly toxic are removed from further consideration.

Level 2 of the screen uses, where possible, cells from the same patient's tumor and determines the potency of the compounds advanced from Level 1. Typically, six serial dilutions (of four- to five-fold each) for each sample are assayed in triplicate, and the 50% inhibitory concentration (IC_{50}) is determined. Test samples with potency in the nanomolar range for pure compounds and at least the microgram per milliliter range for crude extracts are promoted to the third level of the screen.

Level 3 of the screen is designed to determine the cytotoxic selectivity of samples for tumor cells vs. normal cells. Where possible, the same patient's tissues are used. As in Level 2, six serial dilutions (of four- to five-fold each) are assayed in triplicate for each sample. The diluted samples are then added to the tumor cell and normal cell cultures, and the IC_{50} is determined. A "selectivity index" (SI) is determined based on the IC_{50} for normal cells/IC_{50} for tumor cells. Samples with an SI of three or more are advanced to Level 4 of the screen. Additionally, only purified and well-characterized compounds are promoted for further testing.

Before further testing and to confirm that the compounds are cytotoxic rather than merely interfering with the Alamar blue indicator dye, they are re-bioassayed using two other indicator dyes. Calcein-AM is a fluorescent dye that measures changes in cell membrane permeability, an indicator for one of the penultimate steps of cell death. Luciferase measures the amount of adenosine triphosphate (ATP) synthesis in a chemiluminescent assay. For some compounds, cell death was also confirmed by microscopic examination of Papanicolaou-stained cell preparations.[11]

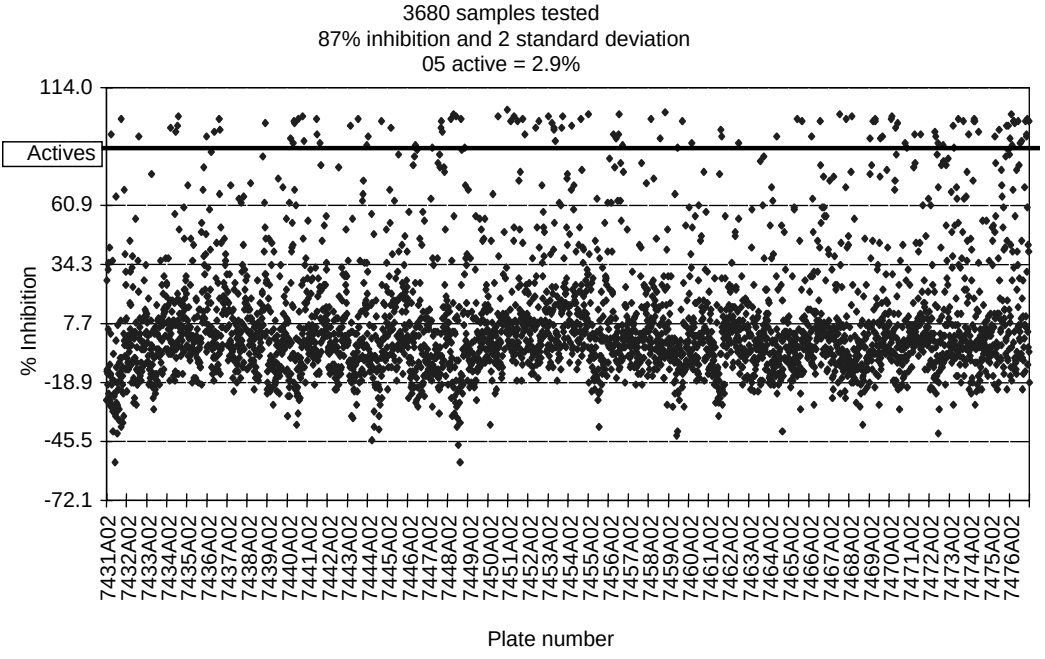


FIGURE 8.4
Sample Cytection run set.

Level 4 of the screen determines the spectrum of activity of the compounds against a wide range of tumor and normal cells derived from different anatomical locations. Tumor cells should include those from patients with different diagnoses and grading, metastatic tumors, and chemoresistant tumors. To properly stratify a compound's *in vitro* antitumor potency and selectivity, a minimum of 25 to 30 tumor specimens from different patients should be tested.

The same dose–response dilution schemes as described for Levels 2 and 3 are used for these assays, and the IC_{50} is determined. The compounds having the widest spectrum of nanomolar (n*M*) antitumor potency and an SI of 3 or greater for a panel of normal cells are considered to be "Hits" in the screen and to merit further evaluation and testing.

The results of an antitumor screen are summarized in Table 8.1. The attrition table summarizes the results from 338,072 samples tested against tumor cells derived from soft tissue sarcomas. Given that the samples included one combinatorial collection with approximately 1.5 million compounds and that each natural product extract most likely contained 100 or more, the total number of compounds tested in this screen exceeded 5 million. As shown in the first column of Table 8.1, the samples were from 11 collections composed of single synthetics, compounds synthesized by combinatorial chemistries, and purified natural products and extracts. The natural products were derived from microorganisms (actinomyces and fungi), plants, and marine invertebrates.

The second column of Table 8.1 lists the number of samples from each collection tested for activity in Level 1 of the screen. The third column shows that 9717 samples were identified in Level 1 of the screen according to the selection criteria (at least 80% inhibition of tumor cells within two to three standard deviations from the mean) and were tested for potency in Level 2 of the screen. This represented 2.9% of the samples originally screened in Level 1.

TABLE 8.1

Attrition of Samples Screened

Collection	Level 1 Activity	Level 2 Potency	Level 3 Selectivity	Level 4 Spectrum	Hits
Single synthetic	10,080	294	35	3	2
Combinatorial[a]	18,160	158	15	1	0
Synthetic/plant	20,080	529	82	3	0
Fungal	21,200	240	29	6	0
Combinatorial	9840	273	5	0	0
Combinatorial	108,320	3503	215	5	0
Actino/fungal	45,200	1612	466	11	0
Marine	2099[b]	277[c]	182	0	0
Single synthetic	100,755	2831	817	52[e]	5
Single synthetic	498[d]	N/A	40	3	0
Plant	1840[d]	N/A	135	0	0
Total	338,072	9717	2021	81	7 classes

[a]Library contains pools of 1.5 million compounds.
[b]880 samples assayed in primary; 1098 samples assayed directly for potency.
[c]Estimated yield due to direct potency testing.
[d]Assayed directly for potency.
[e]Includes analogs.

The fourth column of Table 8.1 indicates that 2021 samples (or 0.6% of the original samples) identified in Level 2 as having n*M* potency for purified compounds or microgram-per-milliliter potency for extracts were tested for selectivity in Level 3. The normal cells used in the screen were, historically, lung cells for two reasons. First, some of the sarcomas first used in the screen were located in the lung and, therefore, normal tissues used in the bioassays were also from that organ. Second, lung tissue contains a varied composition of cells including epithelial cells (ciliary and goblet), stromal cells, and inflammatory cells and, thus, is desirable in determining the cytotoxicity of compounds for a wide range of cell types.

The fifth column of Table 8.1 shows that 81 samples (or 0.024% of the original total) were selected as having an SI of 3 or higher (IC_{50} of normal cells divided by IC_{50} of tumor cells) and were tested for their spectrum of activity in Level 4 of the screen. The high rigor and stringency of this screen are exhibited in the last column. The "hit" compounds emerging from Level 4, those having the widest spectrum of antitumor activity and an SI of 3 or greater for a panel of tumor cells vs. normal cells, belonged to only seven "classes" of compounds. This represents 0.0021% or 1:48,000 of the original samples tested. This "hit rate" becomes higher when the estimate of 5 million compounds screened is considered. This demonstrates that it is only the rare compound that can demonstrate n*M* potency and the desired selectivity for primary tumor cells over primary normal cells in a wide spectrum of cell types.

It is noteworthy that redundancy within and among collections meant that classes of compounds, rather than discrete compounds, emerged from the screen, indicating that compounds with the desired biological activity shared a similar structure. In one instance, the same class of compounds was discovered in two different collections.

Results of the screen also show that compounds that passed four levels of the screen were from single synthetic collections, that combinatorial collections lacked the "chemical diversity" demanded by this type of screen, and that samples tested in Level 4 from natural product collections were very toxic to the liver or kidney.

Table 8.2 depicts the spectrum of activity for compound CP 6741. Cells from 11 tumors (shown vertically) and seven normal tissues (shown horizontally) were chosen for evaluation. Their IC_{50} values are alongside or below the description of the tissue used in the bioassay. Highlighted in light grey are pairs of tumor and normal tissues where the SI is 3 or greater. Pairs highlighted in medium grey represent an SI <3 and ≥ 1, and pairs highlighted in dark grey represent an SI <1.

TABLE 8.2

Compound 6741 Level 4 Spectrum Profile

					Normals			
		Lung	**Ovary**	**Kidney**	**Spleen**	**Liver**	**Heart**	**Blasts**
Tumors	**IC_{50} (μM)**	**>1.27**	**>1.94**	**0.689**	**0.033**	**0.370**	**0.373**	**0.365**
Lung	0.068	>18.70	>28.57	10.15	5.45	5.45	5.49	5.38
Renal cell	>2.07	N/A	N/A	>0.33	>0.02	>0.18	>0.18	>0.18
Ovarian	0.035	>36.60	>55.91	19.86	0.94	10.66	10.75	10.52
Sarcoma	0.042	>29.88	>45.65	16.21	0.77	8.71	8.78	8.59
Melanoma	0.076	>16.67	>25.46	9.04	0.43	4.86	4.90	4.79
Neuroblastoma	0.191	>6.65	>10.16	3.61	0.17	1.94	1.95	1.91
Liver	0.506	>2.51	>3.83	1.36	0.06	0.73	0.74	0.72
Bladder	0.105	>12.10	>18.48	6.56	0.31	3.52	3.55	3.48
Sarcoma	0.064	>19.91	>30.41	10.80	0.51	5.80	5.85	5.72
Mesothelioma	0.08	>15.87	>24.25	8.61	0.41	4.63	4.66	4.56
Pancreas	0.262	>4.85	>7.40	2.63	0.12	1.41	1.42	1.39

As might be expected from a screen in which the target tumor was a soft tissue sarcoma, those tumors were among the most sensitive to the compound, along with ovarian tumors, mesotheliomas, melanomas, and lung carcinomas. Tumors least sensitive to this compound included pancreatic carcinomas, neuroblastomas, and especially, renal cell carcinomas. As a rule, for the seven classes of compounds identified in the screen, soft tissue sarcomas, ovarian carcinomas, and mesotheliomas were the most sensitive tumor types.

8.4 Epilogue

Compounds discovered in this screen are now in various stages of further evaluation or optimization, preclinical testing, and clinical investigation. Their structures and physicochemical properties combined to distinguish them as having drug-like properties and analogs to the "parent" compounds have demonstrated superior biological or physicochemical properties. Results from the pharmacological profiling of some of the compounds and preliminary pharmacology and toxicology testing in animals have been encouraging.

8.5 Discussion

This chapter describes the use of both diseased and normal primary patient cells in high-throughput differential screens that can lead to the discovery of new and more effective antitumor compounds. The underlying Cytection technology allows for the efficient harvesting, isolation, and short-term culturing of well-characterized tumor and normal cells suitable for HTS bioassays. Designed to be rigorous and stringent, the four-level screen is capable of discovering compounds that are potent against tumor cells but comparatively less so against normal cells. The author would encourage other members of the drug discovery community to consider this approach as either an alternative or complement to current technologies.

Acknowledgment

I would like to gratefully acknowledge and thank Candy Cottier for her editorial assistance.

References

1. Fenster, J.M. and Tag Ook, P.E., in *Mavericks, Miracles, and Medicine*, Carroll & Graf Publishers, New York, 2003, pp. 135–148.
2. Stock, C.C. et al., Sarcoma 180 inhibition screening data, *Cancer Res.*, 15(Suppl. 2), 179–331, 1955.

3. Zubrod, C.G. et al., The chemotherapy program of the National Cancer Institute: history, analysis and plans, *Cancer Chemother. Rep.*, 50, 349–350, 1966.
4. Abbott, B.J. et al., Screening data from the Cancer Chemotherapy National Service Center screening laboratories, XXX: Plant extracts, *Cancer Res.*, 26, 34–165, 1966.
5. Smith, C.G., personal communication.
6. Hamburger, A.W. and Salmon, S.E., Primary bioassay of human tumor stem cells, *Science*, 197, 461–463, 1977.
7. Shoemaker, R.H. et al., Application of a human tumor colony-forming assay to new drug screening, *Cancer Res.*, 45, 2145–2153, 1985.
8. Skehan, P. et al., New colorimetric cytotoxicity assay for anticancer-drug screening, *J. Natl. Cancer Inst.*, 82, 1107–1112, 1990.
9. Smith, C., Drug discovery in reverse, *Nature*, 428, 227–228, 2004.
10. Ahmed, S.A., Gogal, R.M. Jr, and Walsh, J.E., A new rapid and simple non-radioactive assay to monitor and determine the proliferation of lymphocytes: an alternative to (3H)thymidine incorporation assay, *J. Immunol. Methods*, 170, 211–224, 1994.
11. Papanicolaou, G.N., A new procedure for staining vaginal smears, *Science*, 95, 438–439, 1942.

9

The Evolving Role of the Caco-2 Cell Model to Estimate Intestinal Absorption Potential and Elucidate Transport Mechanisms

Jibin Li and Ismael J. Hidalgo

CONTENTS

9.1 Introduction

Oral administration is the preferred route for delivering therapeutic agents to the systemic circulation and eventually to the intended pharmacological action site(s), because of convenience and patient compliance. However, to reach the site of action, drug molecules must overcome absorptive and metabolic barriers that restrict drug access to the systemic circulation. The intestinal epithelium constitutes a selective permeability barrier between the blood and the environment. It permits the absorption of nutrients such as, sugars, amino acids, peptides, lipids, and vitamins, while it restricts the access of xenobiotics, digestive enzymes, and bacteria.[1] The intestinal epithelium is composed of a single layer of heterogeneous cells, which include enterocytes or absorptive cells, undifferentiated (crypt) cells, endocrine cells, and goblet cells.

The absorptive cells originate in the villus crypt as undifferentiated cells. As the cells move along the crypt–villus axis they undergo enterocytic differentiation to develop a tall, columnar appearance, and microvilli and tight junctions. The tight junctions join adjacent cells to form a cell monolayer, which serves both as physical and biochemical barrier to intestinal drug absorption. As physical barriers, the tight junctions prevent free transepithelial drug diffusion via the paracellular route. Paracellular permeation represents a minor transepithelial transport pathway and drugs that undergo only paracellular transport achieve low to moderate extent of absorption.[2,3] The transcellular route of permeability is accessible to compounds that display suitable physicochemical properties. Most drugs with high extent of absorption traverse the intestinal epithelial mucosa through the cell membrane.[4,5]

Passive transcellular transport across the intestinal epithelium involves three discrete steps: (1) uptake across the apical membrane, (2) diffusion through the cytoplasm, and (3) efflux across the basolateral membrane. Occasionally, drug molecules without favorable physicochemical properties traverse the intestinal epithelium using endogenous membrane transporters.[6–8] In addition, the intestinal mucosa, with its numerous drug-metabolizing enzymes and efflux transporters, such as P-glycoprotein (Pgp), functions as a biochemical barrier.[9]

Technological advances in combinatorial chemistry and high-throughput screening enable drug discovery scientists to synthesize and identify large numbers of pharmacologically active drug candidates in a short time. Thus, there has been an increasing demand for more rapid techniques to evaluate the pharmaceutical properties (e.g., permeability, metabolism) of drug candidates.[10–12] During the late 1980s and early 1990s, pharmaceutical scientists began to use cell-culture models for drug-transport studies. Among the various cell lines established, Caco-2, derived from human colon adenocarcinoma, gained acceptance because it undergoes spontaneous differentiation under normal culture conditions.[13,14] This differentiation is characterized by the development of features of small intestinal cells such as tight junctions and microvilli.[13–15] In addition to expressing drug-metabolizing enzymes (e.g., aminopeptidases, esterases, sulfatases, and cytochrome P_{450}), they also have uptake transporters responsible for the absorption of bile acids, large neutral amino acids, biotin, monocarboxylic acids and PEPT1, and efflux transporters (e.g., P-glycoprotein, BCRP, MRP2, and MRP3).[16–32]

9.2 Caco-2 Cells as an Intestinal Permeability Model

Early studies with Caco-2 cells followed two main objectives. First, there was a great deal of interest in validating the utility of Caco-2 cells in intestinal absorption studies. The focus of a

number of reports was to compare permeability coefficients obtained in Caco-2 monolayers with drug absorption *in vivo*. These comparisons demonstrated a correlation between *in vitro* permeation (P_{app}) and the extent of absorption (F_{abs}) in humans. The interest in this correlation can be easily understood. A strong correlation between *in vitro* P_{app} and F_{abs} in humans is extremely valuable because it demonstrates the utility of *in vitro* P_{app} values to predict the absorption potential of drug candidates. Much research activity has been dedicated to further the characterization of Caco-2 monolayers as an intestinal permeability model. These efforts have sought to compare various aspects of Caco-2 cells and intestinal enterocytes to validate the utility of Caco-2 cells in the study of mechanisms of drug absorption.

9.2.1 Caco-2 Cell Culture

For transport experiments, cells are generally seeded on polycarbonate filters (Transwell™) uncoated or coated with extracellular attachment factors (e.g., rat tail collagen, type I). The cells are cultured in high-glucose Dulbecco's modified Eagle's medium (DMEM), supplemented with 10% fetal bovine serum (FBS), 1% nonessential amino acids (NEAA), 1% L-glutamine, penicillin (100 U/mL) and streptomycin (100 μg/mL) at 37°C under a humidified air–5% CO_2 atmosphere for about 3 weeks.[15] Prior to starting transport studies, the integrity of cell monolayers is assessed by measuring their transepithelial electrical resistance (TEER) and determining the transepithelial fluxes of markers of passive paracellular diffusion. TEER measurements are quick, easy, and nondestructive. TEER can be measured using an EVOM Epithelial Voltammeter (World Precision Instrument, New Haven, CT, USA). In our laboratory, only TEER values of Caco-2 monolayers ranging from 450 to 650 Ω cm^2 are considered acceptable for permeability studies. Monolayers with TEER values outside this range are discarded.

The permeability coefficients P_{app} of passive paracellular flux markers such as mannitol and Lucifer yellow are more sensitive than TEER values in indicating monolayer integrity.[15,33] Other permeability markers such as inulin and PEG 4000 have also been used, but because of their molecular size they are less sensitive and less relevant than mannitol as indicators of monolayer integrity in drug-permeability studies.[15,34] Owing to the difficulty of detecting mannitol by common analytical techniques, radioactively labeled mannitol is used routinely. When the use of radioactive material is undesirable, Lucifer yellow, a fluorescent passive paracellular permeability marker, represents a good alternative marker of monolayer integrity. The acceptable transepithelial flux rate of Lucifer yellow is less than 4×10^{-7} cm/sec in our laboratory. Because P_{app} value of Lucifer yellow is more sensitive as an indicator of monolayer imperfections, monolayers of Lucifer yellow with P_{app} greater than 4×10^{-7} cm/sec are rejected, even if their TEER values are in the acceptable range.

9.2.2 Caco-2 Permeability Assay

Permeability screening experiments are commonly carried out in 12- or 24-well transwell plates. Cell monolayers grown on porous polycarbonate filters are washed and bathed in a transport buffer such as Hanks' balanced salt solution (HBSS), containing 2 mM glucose. For pH dependence experiments, the transport buffer solution can be adjusted using 10 mM MES for pH 5.5 to 6.5 or using 10 mM HEPES for pH 6.5 to 7.4. For low-solubility compounds the transport solution can be supplemented with 1 to 4% BSA to enhance the apparent solubility. Although organic solvents can be used for this purpose, the concentration of these solvents needed to produce an effective increase in the solubility of most poorly soluble compounds is likely to cause cellular damage. At time zero, compounds

dissolved in transport solution are dosed to either the apical chamber (apical-to-basolateral) transport or the basolateral chamber (basolateral-to-apical) transport (Figure 9.1). Then the transwell plates are incubated at 37°C with appropriate agitation to reduce the thickness of the unstirred water layer. Samples are collected from the receiver chambers at preselected time points (e.g., 30, 60, 90, and 120 min) and replenished with blank transport buffer. To determine the percent recovery (or mass balance), samples should be collected from the donor chamber at the end of the experiment. Drug transport is expressed by apparent permeability coefficient P_{app} as:

$$P_{app} = \frac{\Delta Q}{\Delta t A C_0} \tag{9.1}$$

where ΔQ is the amount of drug transported during the time interval Δt, A the surface area, and C_0 the initial concentration in the donor chamber. P_{app} values are used to determine the absorption potential of the test compounds. In our laboratory, the absorption potential is categorized as follows: high, $P_{appA\text{-}B} > 1 \times 10^{-6}$ cm/sec; medium, 0.5×10^{-6} cm/sec $< P_{appA\text{-}B} < 1 \times 10^{-6}$ cm/sec; and low, $P_{appA\text{-}B} < 0.5 \times 10^{-6}$ cm/sec. For passive diffusion and nonsaturated transporter-mediated transport, P_{app} values are independent of the initial drug concentration. A decrease in $P_{appA\text{-}B}$ value with increasing initial drug concentration indicates saturation of transporter-mediated transport. When bidirectional transport is measured, the ratio of the B-A P_{app} divided by the A-B P_{app} ($P_{appB\text{-}A}/P_{appA\text{-}B}$) can be used to assess the involvement of efflux transport mechanisms (e.g., P-glycoprotein).

Compared to *in vivo* studies, the Caco-2 model substantially increases the speed at which absorption potential can be estimated and reduces the amount of drug substance needed. However, manually performed assays are still too slow and labor intensive compared to biological high-throughput screening assays. Caco-2 cells take about 3 weeks to form monolayers of fully differentiated cells. At this point, Caco-2 monolayers are used to evaluate absorption potential under a variety of permeability protocols. In order to further expedite the process of absorption potential assessment, efforts have been made to increase the throughput of Caco-2 transport experiments.

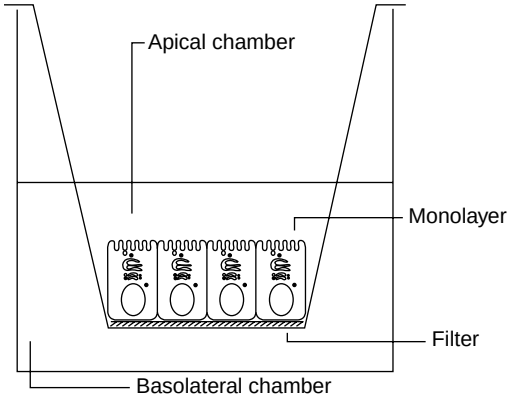


FIGURE 9.1
Schematic representation of the Transwell™ system.

9.2.3 Estimation of Absorption Potential

A good correlation between drug absorption in rats and humans has been demonstrated.[35–39] This correlation permits the use of absorption in animals to predict drug absorption in humans. Ideally, the extent of absorption of new molecular entities should be tested in the lead-optimization state because this information would allow the selection of lead compounds with favorable absorption characteristics. However, the routine use of animals to determine extent of absorption of large numbers of compounds is not feasible for several reasons. First, if the compound undergoes presystemic metabolism, determination of extent of drug absorption requires the administration of radioactively labeled compound, which is rarely available during the lead optimization phase. Second, absorption studies in animals require relatively large amounts of material, drug quantification in plasma is more difficult, and experiments are more laborious and time consuming. Third, *in vivo* studies provide little insight into intestinal absorption mechanisms (e.g., passive diffusion vs. transporter-mediated transport and paracellular vs. transcellular diffusion). *In vitro* systems, which comprise only some of the factors responsible for intestinal drug absorption, have greater utility when applied to the study of mechanisms of drug absorption. Thus, the use of cell-culture models such as Caco-2 constitutes an attractive alternative for estimating the absorption potential of lead compounds. The Caco-2 permeability assay allows the determination of the absorption potential of large numbers of compounds in a short time and with little amount of compound. For instance, Caco-2 experiments normally are conducted in simple buffer solution for a short time (1 to 2 h) with low consumption of test material (submilligram quantities).

9.2.4 Interlaboratory Variability in Caco-2 Permeability Measurements

Reports showing a good correlation between Caco-2 P_{app} and fraction absorbed in humans (F_{abs})[35–39] enhanced the attractiveness of Caco-2 permeability measurements to estimate oral-absorption potential. However, with increasing utilization of Caco-2 cells, a large interlaboratory variability in permeability measurements has become evident. This interlaboratory variability is probably due to biological factors such as culture conditions, passage number, days in culture, and cell viability, and experimental design factors such as composition and pH of transport buffer, drug concentration, use of cosolvents, sampling times, use of mixing conditions, and calculation procedure.

A large interlaboratory variability in the two most common indexes of monolayer integrity, TEER and P_{app} value of mannitol, has been reported.[35–39] This variability in P_{app} values does not apply only to mannitol because when Caco-2 P_{app} data from five laboratories were plotted against known F_{abs} in humans, the interlaboratory variability was extremely large.[40] This variability in P_{app} values makes it extremely difficult to combine, or even compare, data from different laboratories. In an attempt to address the need for standardization of Caco-2 permeability measurements, U.S. FDA scientists, conducted a thorough examination of the complex issues involved in this type of measurements, and concluded that standardization of the assay would be difficult and could hinder further improvements in the area.[41] They recommended that a more general approach to permeability measurements be followed. This approach, described in the context of the Biopharmaceutics Classification System (BCS), permits the use of alternative cell lines in permeability measurements, as long as their performance is appropriately validated.[41] Validation can be accomplished by demonstrating a correlation between *in vitro* permeation and fraction absorbed (F_{abs}) in humans for a set of about 20 compounds. The set contains compounds whose extent of absorption is low ($< 50\%$), moderate (50 to 89%), and high (at least 90%).[41] As validation requires correct rank ordering but not an absolute value

for any of the reference compounds, the interlaboratory variability problem is unlikely to be resolved in the near future. Moreover, the use of alternative cell lines does not necessarily avoid these issues as other cell lines might have unique drawbacks which are likely to be noticed only after widespread use. In any case, performing all the experiments in a single laboratory using robust protocols eliminates interlaboratory variability and should increase the reliability of the permeability coefficients values.

9.3 Use of Caco-2 Cells in Drug Discovery

9.3.1 High-Throughput Permeability Screening

Partly due to the limited throughput of Caco-2 permeability measurements, the structure–activity evaluation of compounds tested in the hit-to-lead phase is done with minimal permeability information, at best. Given the importance of membrane permeability in drug absorption, early consideration of the permeability characteristics of hit compounds would enhance the drug-like quality, and ultimately the probability of success, of selected lead candidates. To incorporate permeability information into the hit-to-lead phase of the drug discovery process it is necessary that permeability measurements be made quickly and with small amounts of material. Thus, efforts have been made to automate and miniaturize the Caco-2 permeability assay.

One approach used to facilitate the automation of Caco-2 permeability measurements was the development of specially designed liquid-handling machines. The gradual implementation of automated permeability assays required a continuous change in the type and size of transport plates. While the early work with Caco-2 cells was done in 6-well transwell trays,[15,17,21–24,34,42] the need for higher throughput was quickly accompanied by a shift toward 12- and 24-well plates. Unlike the standard 24-well plates, where each insert is independent, the 24-well HTS plates were designed for use with liquid handlers (24-well HTS) and consist of a single piece with 24 inserts built-in. One important pitfall of the single-piece 24-well HTS is the difficulty in dealing with bad monolayers. For example, in standard transport plates, inserts containing monolayers with low transepithelial electrical resistance can be excluded from the experiment and replaced with suitable monolayers from other plates. However, in the single-piece 24-well HTS system bad inserts cannot be discarded. Their exclusion from the experiment can be extremely difficult. Often, all the monolayers are dosed but only the samples from the good monolayers are processed, a procedure that creates a lot of unnecessary, and potentially confusing data.

To further exploit the benefits of automation, permeability plates have undergone noticeable miniaturization, as evidenced by the recent development of a single-piece 96-well (HTS) transwell plate. The successful implementation of 96-well HTS plates in permeability assays is not necessarily trivial because of the difficulty associated with achieving homogeneous seeding of such small filters. Nonhomogeneous seeding could lead to cell stacking,[43] which would affect the accuracy and reproducibility of permeability measurements. The difficulty related to the treatment of bad monolayers, mentioned in connection with the single-piece 24-well HTS plates, will probably be enhanced in the 96-well HTS plates.

Despite this concern, it is likely that the use of the 96-well HTS plates will increase substantially in the near future because they allow a potential fourfold throughput increase in the permeability assay compared with 24-well HTS plates.[44,45] The 96-well

HTS plates permit to determine drug permeability across a cell monolayer with a throughput similar to that of the parallel artificial membrane permeation assay (PAMPA), which measures rate of diffusion across a lipid layer.[46] As is the case with PAMPA, the tiny surface area of the filters of the 96-well HTS presents an analytical challenge for compounds with low-to-moderate permeability.

Generally, automation formats focus on determining A-B P_{app} values at the expense of B-A P_{app} values. In an attempt to facilitate the automation of B-A P_{app} measurements, Garbeg et al.[47] described a procedure that involved culturing Caco-2 cells on either the inside or outside surface of transwell insert filters. When cells are cultured on the inside surface of filters (standard procedure) the apical domain of the monolayers faces up and when cells are cultured on the outside surface of the filters the apical domain of the monolayers faces down (the well). They found that cells grown on the outside surface of the filters achieved barrier properties similar to those of cells grown on the inside surface of the filters. However, the former failed to express appreciable levels of the efflux transporter, P-glycoprotein, and the uptake transporter for large neutral amino acids.[47] This approach, published in 1999,[47] has not found wide acceptance.

Miniaturization and automation have resulted in a dramatic increase in the throughput of permeability measurements. This increase would be of little value without a concurrent increase in the analytical methodology used to quantify the samples generated in permeability assays. LC/MS(MS) technology, together with automation and miniaturization, has enabled the overall increase in throughput of the permeability assay.[48] LC/MS(MS) not only makes method development faster and easier compared to HPLC, but also accelerates sample analysis. The specificity of MS/MS permits the analysis of multiple analytes, originating from either cassette dosing or postpermeability sample pooling, in a single injection.[48,49]

9.3.2 Cassette Dosing

The preferred way to determine permeability is applying a dosing solution containing one compound (i.e., discrete dosing), and taking samples from the receiver chamber to calculate P_{app}. This type of dosing makes the data easy to interpret. However, to increase throughput, compounds are also tested applying a mixture (i.e., cassette dosing) to the donor chamber. Cassette dosing has been used in pharmacokinetic studies with varying degrees of success.[50,51] It relies on the power of LC/MS/MS to quantify the different compounds in the sample. Although this approach reduces the number of cell monolayers, experimental time, and number of injections, it has inherent risks. For example, in pharmacokinetics studies the area under the plasma concentration vs. time curve (AUC) of some compounds is often higher after cassette dosing than after discrete dosing,[52] a situation that might result in a large number of false-positive results. Cassette dosing could cause saturation of enzymes involved in presystemic elimination and intestinal or hepatic transporters. Permeability studies with Caco-2 cells dosed with cassettes of compounds found comparable P_{app} values for four- and eight-compound cassettes compared with discrete dosing.[49] The success of the cassette approach depends on the careful selection of the individual compounds in the cocktail. By choosing compounds that are known not to interact with transporters or enzymes, it is possible to evaluate the utility of cassette dosing under ideal conditions. As is the case in pharmacokinetics, cassette dosing might work well for the right combination of compounds. Unfortunately, in a real discovery program the transport and metabolism characteristics of test compounds are often unknown. Caco-2 cells express a large number of intestinal transporters and enzymes.[17,19–32] Thus, the potential for drug–drug interactions when testing cassettes of compounds of unknown properties should not be ignored. Considering these risks, it

appears reasonable to think that the use of cassette dosing in permeability assays should be followed by a secondary discrete screening step to try to minimize the number of false positives. For example, any compound exceeding a given threshold following cassette dosing should be tested individually to determine whether the results from cassette dosing are real or a false positive due to drug–drug interaction(s). A disadvantage of this technique is that more pre-experiment planning is required to ensure analytical compatibility among the compounds in the cassette. Mixing compounds with different intrinsic permeability characteristics may result in unsuitable sampling times, sampling volumes, or compounds concentration in the receiver chamber. For example, for high-permeability compounds the experiments should have a shorter duration and the sampling intervals times should be shorter, compared with low permeability compounds.

9.3.3 Accelerated Cultures

Caco-2 cells seeded at a density of approximately 60,000 cells/cm^2 and cultured under standard conditions reach a monolayer density in about 1 week and suitable barrier properties in about 2 weeks.[15,34,42] However, the expression of enzymes and membrane transporters may take longer, and, in some cases, up to 4 weeks are needed for optimal transporter expression.[17,21,53] So, in general, Caco-2 cells are cultured about 3 weeks before being used in permeability studies. To shorten this culture time, several modified Caco-2 culture protocols have been developed. One of these, termed 3-day Caco-2, consists of seeding the cells at a high density (e.g., 480,000 cells cm^{-2}) and culturing them with Biocoat® media, which contains hormones, growth factors (to ensure faster growth), and sodium butyrate (to induce differentiation). The 3-day Caco-2 monolayers have lower TEER values and lower expression of Pgp than the 3-week Caco-2 monolayers.[54] Due to the lack of evidence for the expression of transporters in the 3-day Caco-2 cells, this system is proposed as a tool to screen for permeability of compounds that undergo passive diffusion, but not for compounds that might be substrates for membrane transporters. Also, the lower TEER value of the 3-day Caco-2 monolayers has been argued to be more representative of the intestinal permeability, which is lower than that of the average 3-week Caco-2 monolayers.[54] For passively transported compounds, there was generally a good correlation between 3-day and 3-week Caco-2 monolayers, but the 3-day monolayers tended to overpredict P_{app} when compared with the 3-week monolayers.[54] Another concern with the 3-day Caco-2 cells is that, the higher seeding densities used with this approach, are likely to increase the likelihood of cell overgrowth (cell stacking) compared to the 3-week Caco-2 cells, which are seeded at lower density. Other protocols for accelerated Caco-2 cell culture have also been developed.[55–57] In general, Caco-2 cells cultured in accelerated protocols show lower levels of Pgp and uptake transporters,[55–57] with the exception of one study that found comparable levels of Pgp and PepT1 between modified 3-day Caco-2 cells culture and a 3-week Caco-2 cells.[57]

9.4 Use of Caco-2 Cells in Drug Development

9.4.1 Biopharmaceutics Classification System

The publication of the U.S. FDA Guidelines allowing the waiver of bioequivalence study for compounds classified as Class I based on the Biopharmaceutics Classification System

has bolstered the role of Caco-2 cells in permeability measurements.[58] These guidelines make it possible to avoid costly and time-consuming bioequivalence studies for a small number of compounds that meet the Class I requirements. The FDA guidelines did not identify either Caco-2 or any other specific cell line for BCS permeability classification. Although in principle any cell line can be used as long as it is validated according to criteria described in the guidelines, Caco-2 cells have become the preferred cell line for this type of test. The primary reason for this preference is that many pharmaceutical scientists are familiar with Caco-2 cells and they are usually available in most pharmaceutical laboratories. Since the validation involves a fair amount of work, it does make sense to use a cell line that is widely used throughout the industry rather than undertaking the validation of a lesser known cell line for this limited application.

To meet the permeability requirements for BCS Class I, a compound must be tested in a laboratory that has shown method suitability using Caco-2 cells (or other cell line). After method suitability has been established, the classification of test compounds is done by comparing the permeability coefficient of the test compound to that of a high-permeability reference compound (e.g., pindolol, antipyrine, labetalol). The high-permeability reference compound must have an extent of absorption in human of at least 90%. If the permeability coefficient of the test compound is higher than that of the high-permeability reference compound, the permeability class of the test compound is classified as high. Together with the high permeability reference control the permeability assay must include a low-permeability control (e.g., mannitol, atenolol). The purpose of the inclusion of a low permeability control is to ensure cell monolayer integrity. Leaky monolayers would yield artificially high-permeability coefficient values for the test compound.

The utilization of *in vitro* models in general, and Caco-2 cells in particular, to determine the BCS class of compounds is likely to increase as scientists become more familiar with the FDA guidelines and start to benefit from their application. It is clearly understood that pharmaceutical development efforts are more likely to succeed if the new molecular entity exhibits suitable physicochemical (i.e., drug-like) properties. In addition, the impact of the BCS is expected to increase as the potential for obtaining biowaivers is expanded to include compounds in classes other than class 1. For example, two recent publications propose the extension of the BCS guidelines to make some class 2 compounds eligible for a BCS biowaver. It has been suggested that, acidic drugs exhibit class 1 characteristics at $pH > 5$, which is the relevant pH for intestinal drug absorption, and high-solubility requirements should only apply at $pH > 5$.[59] This proposal would make weakly acidic drugs (e.g., nonsteroidal anti-inflammatory drugs), currently classified as class 2, eligible for biowavers. In contrast, an AAPS/FDA sponsored workshop concluded that, despite the industry's interest in extending biowaver to class 2 compounds, current scientific understanding of the role of low solubility on drug absorption and product performance is insufficient to assess the risk of low solubility on product failure.[60]

9.4.2 Permeability Enhancers

Two factors that often limit intestinal absorption are poor aqueous solubility and low membrane permeability. When the reason for low absorption is insufficient membrane permeability there are a few experiments that can be undertaken to enhance permeability. Several approaches have been evaluated in Caco-2 cells. First, the use of prodrugs to increase hydrophobicity results in higher affinity for the cell membrane. Second, the synthesis of prodrugs that target intestinal mucosal transporters such as PepT1 has been shown to increase systemic drug absorption.[61–63] Third, drug permeability can be enhanced by modifying the barrier properties of the cell layer through opening of intercellular junctions or

increasing membrane fluidity. Although these approaches to increase drug absorption have been tested in animals and intestinal tissue *in vitro*.[64–66] Caco-2 cells have numerous advantages in this type of studies. These advantages include: (1) greater experimental flexibility (e.g., pH, time frame, directionality), (2) ease of data interpretation, (3) samples are cleaner and thus drug analysis is easier compared to plasma, (4) requires relatively small amounts of material, (5) greater sensitivity to experimental variables permit to screen large numbers of potential enhancers, and (6) experiments are simple and rapid.

The flexibility of the Caco-2 cell permeability model makes it possible to evaluate the relative contribution of specific permeability pathways. The permeability pathway of clodronate and other biphosphonates is passive, paracellular diffusion.[67] The low permeability of these compounds is due to the restrictive diffusion through the intercellular junctions and to the formation of complexes with calcium. Thus, a sound strategy to increase the permeation of this type of compounds is to shift the permeability pathway from paracellular diffusion to transcellular diffusion. This was accomplished by Raiman et al.,[67] who synthesized a series of ester prodrugs of clodronate having varying degrees of lipophilicity. After determining the permeability coefficients in Caco-2 cells of these prodrugs, they found that, when log D reached 1, the permeability pathway of the esters of clodronate shifted from paracellular to transcellular.[67] This type of mechanistic evaluation of intestinal permeability mechanisms is difficult to obtain in more complex systems.

However, when the low absorption is due to low membrane permeability, available options to increase absorption are very limited. To increase the absorption of compounds with low permeability it is necessary to alter the barrier properties of the intestinal mucosa through permeability enhancers. Permeability enhancers can act by either opening tight junctions or by modifying the cell membrane. In principle, lowering the diffusional resistance of the intestinal mucosa might potentially allow the penetration of foreign substances into the systemic circulation, which could result in infections and toxicity. Therefore, it is highly desirable to conduct studies that make it possible to fully understand the interaction of any permeability enhancer with the intestinal mucosa. Questions such as, onset of permeability enhancement effect, reversibility of the enhancement, size limit for "bystander" molecules permeation, mucosal damage, and fate of the enhancer, should be addressed. Very few systems have the necessary versatility to be useful in answering this type of questions. In one study with Caco-2 cells, which have found great acceptance in this type of studies, the role of nine common excipients on the transepithelial permeation of seven low-permeability compounds was evaluated.[68] Most excipients used in this study had no influence on the ability of low-permeability drugs to permeate Caco-2 cell monolayers. The clear exception was SLS, which increased the permeation of all the drugs tested.[68] Other excipients had limited effect on Caco-2 permeability. For example, Docusate sodium caused a moderate increase in cimetidine permeability and Tween 80 had marked increase on the permeability of furosemide and cimetidine, two known efflux-transporter substrates.[68] Although these studies provide useful insight into the effect of excipients on different transcellular transport pathways, the evaluation of potential detrimental effect on the cells was very limited for two reasons. First, in general, excipients were tested at only one (usually low) concentration, which makes it difficult to evaluate the true potential for a detrimental effect on the cell monolayers. Also, because the concentrations of excipients were different, it is not possible to make a direct comparison of the relative potential of different excipients to cause cellular damage. Second, evaluation of monolayer damage was limited to monitoring the impact on the passive paracellular transport pathway and did not assess the effect on cellular biochemistry or viability was obtained.

Thus, a systematic evaluation of excipients focusing on the concentration dependence of effect on barrier properties, biochemical integrity, and cell viability is still needed.

9.4.3 Low-Solubility Compounds

A large proportion of discovery compounds have low aqueous solubility. When intestinal permeability is due to low solubility the use of formulation vehicles that increase the apparent solubility of the compound might increase permeability. Caco-2 cells have been used to evaluate pharmaceutical excipients that increase intestinal permeability through an increase in solubility. Solubility enhancers, which have been used to help solubilize this type of compounds, include cosolvents (DMSO and dimethyl acetamide) or other adjuvants (cyclodextrins, vitamin E TPGS, and albumin). Although concentrations of adjuvants tolerated by Caco-2 cells have been identified,[68,69] evidence supporting "safe" concentration of adjuvants is often controversial. For example, surprisingly few studies have reported on the tolerance to DMSO, the most commonly used solubilizing cosolvent, by Caco-2. Some studies have found that the presence of DMSO causes cellular damage secondary to changes in osmolality.[70,71] Furthermore, it has been reported that DMSO concentrations above 1% have a detrimental effect on the barrier properties of Caco-2 monolayers,[72] and that the presence of 4% HAS was able to prevent the detrimental effect of DMSO concentrations as high as 10% on Caco-2 cells.[72] However, the damaging effect of DMSO on Caco-2 cells is not generally observed. A study with a Caco-2 cell subclone (TC7) showed that 10% DMSO had no effect on the barrier properties of Caco-2/TC7 cells as judged by mannitol permeation and LDH release. Effects of DMSO were observed when the concentration of DMSO exceeded 10%.[73]

Another study found that 5% Tween 80 and HP-β-CD were effective as solubilizing compounds with poor aqueous solubility without causing an increase in the transepithelial permeability of the monolayer integrity marker mannitol. The effect of these excipients on the biochemical integrity of the cells was not reported.

9.5 Mechanistic Studies with Caco-2 Cells

In addition to screening molecules for intestinal absorption, Caco-2 cells have also been used to study mechanisms of drug transport. For many compounds, intestinal permeation involves a transporter to either aid or limit transepithelial transport. The value of Caco-2 cells in this type of studies is due to the fact that these cells express various membrane transporters relevant to drug absorption.[17,19–23,28,30] However, when interpreting results of studies that involve carrier-mediated transport, discretion, and scaling factors may be required because of the difference in expression level of transporters between *in vitro* and *in vivo* systems.[12] Another important consideration in carrier-mediated transport studies is that some transport systems in Caco-2 cells may achieve maximal expression level at different days in culture.[17,21,38,74] Thus, validation of Caco-2 cells for mechanistic studies should include the identification of the time for optimal expression of transporters as well as the qualitative evaluation of the transporters to establish that they are representative of the native intestinal transporters.

9.5.1 Permeability Pathways

Caco-2 cells cultured on filters achieve a monolayer density, and exhibit morphological characteristics similar to enterocytes such as tight intercellular junctions and highly developed microvilli.[13–15,42] Transport of drugs across Caco-2 monolayers is limited by the action of biochemical and physical barriers (Figure 9.2). The biochemical component comprises drug-metabolizing enzymes, uptake transporters, and efflux transporters.[17,19–23,28,30] The

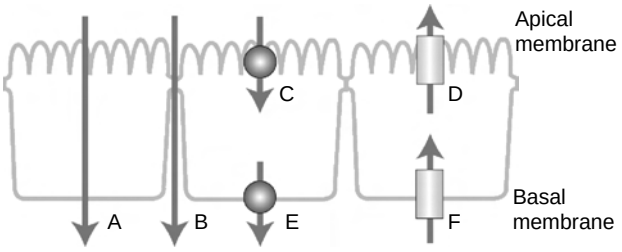


FIGURE 9.2
Major transport pathways in Caco-2 monolayers. A: Passive transcellular; B: Passive paracellular; C: Transporter-mediated apical uptake; D: Transporter-mediated apical efflux; E: Transporter-mediated basolateral efflux; F: Transporter-mediated basolateral uptake.

physical barrier consists of the cell membrane and intercellular junctions. Together, these components bestow on Caco-2 cells monolayers permeability resistance characteristics reminiscent of those found in the intestinal epithelium.

This permeability barrier shows selectivity in that small hydrophobic molecules can partition into and diffuse across the lipid bilayer of the cell membrane, whereas small hydrophilic molecules can only diffuse between cells (i.e., through the intercellular junctions). In addition, the presence of uptake and efflux transporters complicates our ability to predict intestinal permeability based on physicochemical properties alone because transporters may increase or decrease absorptive flux. The complexity of the permeability process makes it difficult to elucidate permeability pathways in complex biological model systems such as animals and tissues. For this reason, cultured cells in general, and Caco-2 cells in particular, have been used extensively to investigate the role of specific permeability pathways in drug absorption.

9.5.1.1 Passive Paracellular Transport

The paracellular transport pathway is an aqueous, extracellular route that traverses the epithelial layer. Paracellular transport generally takes place via passive diffusion and is accessible to small (i.e., molecular weight [MW] < 200) hydrophilic molecules.[75–77] The driving force for passive paracellular diffusion is the electrochemical potential gradient resulting from the differences in concentration, electrical potential, and hydrostatic pressure between the two sides of the epithelial membrane. The barrier restricting the passive movement of solutes through the paracellular pathway is the tight junction.

Although the contribution of the paracellular pathway to drug absorption plays a minor role for high-absorption compound, the converse is true for low-absorption compounds. To characterize the contribution to permeation across Caco-2 monolayers of the paracellular pathway the permeability resistance of this pathway can be modulated by the addition of a calcium-chelating agents such as EGTA or EDTA, which causes opening of tight junctions.[78] Usually, the extracellular concentrations of EDTA or EGTA must be at least 0.5 mM to open the paracellular transport pathway.[78–80] For example, 1.26 mM EDTA increased the permeation across Caco-2 monolayers of ranitidine by 5- to 20-fold without affecting the permeation of ondansetron.[81] This indicates that transport pathways for ranitidine and ondansetron are paracellular and transcellular, respectively.[81]

However, caution must be exercised in the use of Caco-2 monolayers to study paracellular transport. Tanaka et al.[82] found that the TEER values of Caco-2 monolayers (468 Ω cm^2) were higher than those of rat jejunum (37 Ω cm^2) and colon (83 Ω cm^2). Consistent with higher TEER values, the P_{app} of the paracellular permeability marker FITC–dextran (MW = 4000) in Caco-2 monolayers, rat jejunum, and rat colon were 0.22, 2, and 0.95×10^{-5} cm/min,

respectively. That Caco-2 monolayers exhibited the most dramatic decrease in TEER in the presence of 10 mM EDTA indicates that Caco-2 monolayers are more sensitive to tight junction modulation than rat jejunum and rat colon.[82]

9.5.1.2 Passive Transcellular Transport

Passive transcellular transport involves the translocation of solute molecules across the lipid bilayers of the enterocytes. For hydrophobic compounds, transcellular transport is the main route of intestinal epithelial permeation. The importance of this route of permeation is due to its magnitude. The surface area of the cell membrane, which constitutes the transcellular route, accounts for the vast majority (i.e., 99.9%) of the epithelial surface area. The surface area of the tight junctions, which constitutes the paracellular route, represents only 0.01% of the total surface area of the intestinal epithelial layer.[83,84] Generally, the extent of absorption of compounds whose permeability is limited to the paracellular pathway is low and the extent of absorption of compounds that readily traverse the cell membrane is high.[28] However, the involvement of uptake or efflux transporters can change this pattern, making it virtually impossible to predict drug absorption in the absence of specific information on drug–transporter interaction. For example, the absorption of some hydrophilic molecules such as cephalosporins and angiotensin-converting enzyme (ACE) inhibitors is much higher than expected from their intrinsic membrane permeation characteristics.[85,86]

9.5.2 Membrane Transporters Involved in Uptake or Efflux

Caco-2 cells express a number of intestinal epithelial transporters (Figure 9.3). Some of these transporters, which are often referred to as uptake transporters, mediate solute transport from the intestinal lumen into the enterocytes and thus help the drug absorption process. Uptake transporters found in Caco-2 cells include the following: large neutral amino acid transporter, bile acid transporter, oligopeptide transporter (PepT1), monocarboxylic acid transporter, nucleoside transporter, and fatty-acid transporter.[21–23,87,88] Caco-2 cells also express a number of transporters that mediate the apical efflux of solutes thus countering the drug-absorption process. These transporters include members of the ABC superfamily of transporters such as P-glycoprotein (Pgp), MRP, and BCRP. The expression

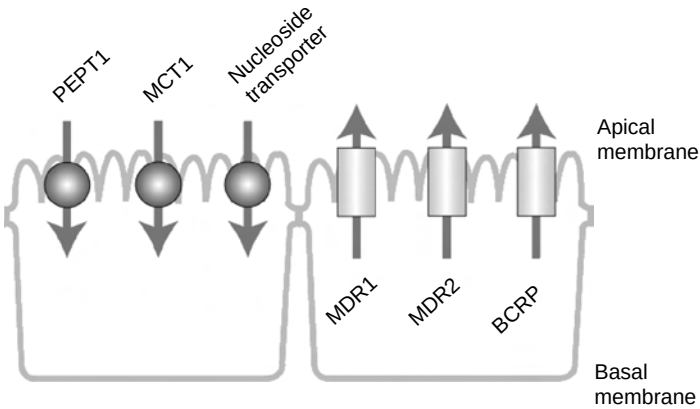


FIGURE 9.3
Partial representation of uptake and efflux transporters expressed in Caco-2 cells.

of uptake and efflux transporters by Caco-2 cells has increased the utility of this cell line in the study of mechanistic aspects of intestinal drug transport.

Before engaging in the use of Caco-2 cells to investigate the role of transporters in intestinal permeability an important consideration that must be kept in mind is the variability of transporter expression. The optimal time to evaluate transporter function is not necessarily correlated with the optimal time for monolayer barrier functions. Although some studies have shown that based on paracellular permeability and TEER values, two indicators of monolayer integrity, the barrier properties are achieved within 2 weeks post seeding; however, the maximal expression of several transporters is not necessarily optimal at this time.[21,53,89] Moreover, the level of transporter expression is influenced by the culture conditions used. Thus, it is unlikely that the level of transporter expression can be standardized based on passage number or post seeding time alone. The reason for this situation is that Caco-2 cell cultures consist of a series of cell subpopulations rather than a single cell population. The existence of subpopulations has made it possible to isolate clones that differ from the general cell culture in the level of expression of characteristics of interest. For example, clones that express high levels of the brush-border marker enzyme, sucrase–isomaltase, bile acid transporter, or Pgp have been isolated.[90–92] The experimental conditions used to culture the cells can affect transporter expression in an undefined manner.

Intestinal transporters are responsible for the high absorption of some drugs whose passive absorption is expected to be low. PepT1 is involved in the absorption of the ACE inhibitors (e.g., enalapril and captopril), penicillins (e.g., amoxicillin) and cephalosporins (e.g., cephalexin, cephradine, and cefaclor). Attempting to exploit the apparently broad substrate specificity of PepT1 conjugates of amino acids and drugs have been made to target this transporter. Ideally, the conjugate which needs to be stable in the gastrointestinal tract, would reach PepT1, undergo translocation across the enterocyte, and would then release the drug in the plasma. If the conjugate is biologically active the release of the drug in plasma is not a requirement. For example, the bioavailability of L-Dopa, which is transported by the large neutral amino acid transporter, is low due to poor membrane permeability and metabolic instability. Following conjugation to L-Phe, the L-Dopa–L-Phe conjugate was recognized by PepT1 resulting in a much higher amount of L-Dopa being transported across Caco-2 monolayers as part of the L-Dopa-L–Phe conjugate.[93] Another example of a successful application of this approach was the conjugation of the antiviral agent, acyclovir. The oral bioavailability of acyclovir is limited (~20%) due to low intestinal mucosal permeation.[61,94] Synthesis of the valine–acyclovir (Valacyclovir) prodrug caused a three- to fourfold increased in acyclovir bioavailability compared to acyclovir.[95]

Evidence of the expression of several efflux transporters in Caco-2 cells is varied. It ranges from quantification of mRNA using real-time PCR to assessment of functionality by determining the vectorial transport of substrates. Although mRNA evidence for MDR1, MRP1, MRP2, MRP3, MRP4, MRP5, BCRP, and LRP, have been found in Caco-2 cells, functionality data have been extensively demonstrated only for MDR1.[20,96] Data showing the functionality in Caco-2 cells of MRP2 and BCRP is very limited and for other efflux transporters is practically nonexistent.

9.5.3 Drug-Metabolizing Enzymes

Caco-2 cells express several intestinal drug-metabolizing enzymes such as peptidases, CYP1A1, CYP1A2, UDP glucuronyltransferases, and phenol sulfotransferases.[26,27,97–100] However, the utility of Caco-2 cells in the study of intestinal drug metabolism has not been fully established. One difficulty is that the level of expression of CYPs, the most abundant intestinal drug-metabolizing family of enzymes, appears low and variable.[96–98] Even

CYP3A4, an isozyme abundant in the intestinal mucosa, is poorly expressed in Caco-2 cells. Two approaches have been used to increase the utility of Caco-2 cells in CYP3A4 metabolism. The first consists of inducing the enzymes by culturing the cells with Di OH vitamin D_3.[101] Based on the scarce number of studies using this system it appears that the application of Di OH vitamin D_3-induced CYP3A4 to drug metabolism studies is very limited. As is the case with transporters, the expression of enzymes is very variable, and the response to Di OH vitamin D_3 treatment is likely to show even greater variability in enzyme activity.

In another attempt to create Caco-2 cells expressing CYP3A4 activity, Caco-2 cells were transfected with human CYP3A4 cDNA.[102] Although some success was reported regarding the expression of CYP3A4 activity, the transfected cells failed to develop stable barrier properties.[102] The absence of recent publications reporting the use of these cells suggests that these cells have failed to receive widespread acceptance. Although the mentioned disadvantages that resulted from the transfection are probably responsible, at least in part, for the lack of acceptance of these transfected cells, it is possible that an interest in simplifying the cells may have played a role. Owing to the difficulty associated with the interpretation of permeability results, especially when membrane transporters are involved, the addition of a metabolism component makes data interpretation even more difficult. This mindset has led to a migration toward simpler permeability models (e.g., PAMPA) and to the use of alternative cell lines such as MDCK.[103,104] It is likely that the future role for Caco-2 cells will be more focused on mechanistic permeability studies and less involved in drug metabolism.

Since 1995 there has been a growing interest in the interplay between metabolism (i.e., CYP3A4) and efflux transport (i.e., Pgp).[105–107] Conceptually, Pgp controls intestinal CYP3A4-catalyzed metabolism by increasing the exposure to CYP3A4 through repeated recycling (i.e., uptake followed by efflux).[107,108] However, the value of Caco-2 monolayers in these studies has been limited because although these cells exhibit a robust expression of Pgp they express very low levels of CYP3A.[109] Although one study successfully induced CYP3A4 with Di OH vitamin D3 and was able to identify the relative roles of Pgp and CYP3A4 on the apical secretion of indinavir metabolites in Caco-2 monolayers,[108] in general, other systems are being increasingly used to evaluate the functional CYP3A–Pgp relationship.[110,111]

9.6 Inappropriate Use of Caco-2 Cells

9.6.1 Formulation Evaluation

Caco-2 cells have been valuable in the estimation of drug absorption potential, transport mechanisms, and effect of permeation enhancers on transepithelial transport.[35,39,53,67–69,78–81] Owing to the sensitivity of the cells and the limited solubility of new molecular entities, Caco-2 permeability studies are routinely done with relatively low concentration of compounds. One way to increase the solubility of these compounds is to use organic solvents. The low tolerability of Caco-2 cells to organic solvents limits the use of this approach in permeability studies.

Although the potential to use Caco-2 cells to screen large numbers of formulations under carefully controlled experimental conditions appears attractive, the assumption that this model is suitable to evaluate drug formulations should not be made. In fact, the utility of Caco-2 cells in formulation evaluation is limited because these cells are sensitive to pharmaceutical excipients.[112] In addition, the dilution of formulations into the simple buffer solutions used in permeability studies is likely to break the physical integrity of the

formulations. Thus, unless this process mimics what happens in the intestinal lumen *in vivo* it is impossible to predict whether the formulation screening results obtained in Caco-2 cells have any relevance to the *in vivo* performance of the formulations.

9.6.2 Bioavailability Prediction

The two major determinants of oral bioavailability are intestinal absorption and hepatic first-pass elimination. Caco-2 cells are useful to predict intestinal absorption. The validity of this application has been demonstrated in a number of studies in which percent drug absorption in humans was correlated with Caco-2 permeability coefficients.[35–39,113,114]

Unfortunately, and probably as a result of a lack of appreciation for the limitations of the model, attempts to use Caco-2 cells to predict oral bioavailability in humans have been reported.[115] Although Caco-2 permeability could predict bioavailability when compounds do not undergo presystemic hepatic metabolism or excretion in the bile, in the absence of detailed information on the pre-systemic handling of drugs it is not possible to anticipate whether hepatic extraction is likely to play a role in limiting drug bioavailability. Taking into account the role of hepatic metabolism in systemic drug availability, a recent study used an integrated Caco-2–hepatocyte system to predict oral bioavailability with reasonable degree of success.[116] However, in general, even if a compound is metabolically stable, it is not possible to assume that Caco-2 permeability will predict bioavailability because of the potential for interaction with hepatic transporters. This concern is due to large number of drugs that are substrates of transporters and thus undergo hepatic uptake or biliary excretion. Under unique circumstances a correlation could be found between Caco-2 permeability and fractions absorbed; however, this situation will probably be the exception rather than the rule. In general, attempts to predict bioavailability based on Caco-2 permeability data, even incorporating *in vitro* metabolic stability data, need to consider the role of intestinal and hepatic transporters to avoid the possibility of generating misleading data.

9.7 Artificial Membrane Methods Used to Estimate Drug Absorption

9.7.1 Parallel Artificial Membrane Permeation Assay

The successful application of *in vitro* models of intestinal drug absorption depends on the ability of the *in vitro* model to mimic the relevant characteristics of the *in vivo* biological barrier. Most compounds are absorbed by passive transcellular diffusion. To undergo transcellular transport a molecule must cross the lipid bilayer of the apical and basolateral cell membranes. In recent years, there has been a widespread acceptance of a technique, artificial membrane permeation assay (PAMPA), to estimate intestinal permeability.[117,118] The principle of the PAMPA is that, diffusion across a lipid layer, mimics transepithelial permeation. Experiments are conducted by applying a drug solution on top of a lipid layer covering a filter that separates top (donor) and bottom (receiver) chambers. The rate of drug appearance in the bottom wells should reflect the diffusion across the lipid layer, and by extrapolation, across the epithelial cell layer.

The advantages of PAMPA over Caco-2 permeability measurements are higher throughput, lower cost, and shorter planning as there is no cell culture involved. PAMPA measurements help identify compounds that have high passive transcellular permeability and

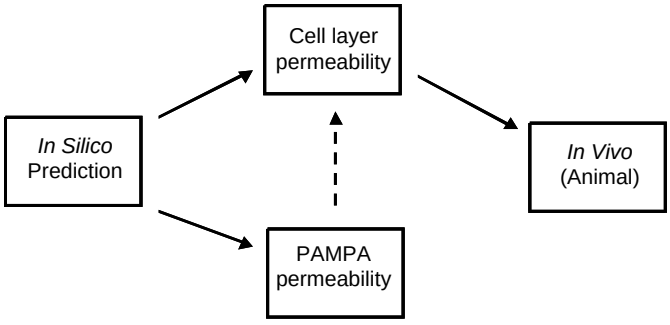


FIGURE 9.4
Complementary use of PAMPA and Caco-2 cells for evaluation of absorption potential. PAMPA measurements are used to discard compounds with clear absorption problems whereas Caco-2 cells would be used to evaluate mechanisms of permeation or reasons for low permeation. It is highly unlikely that PAMPA measurements would be used to select compounds to be tested *in vivo*.

ignore the possible contribution to transepithelial transport of the paracellular pathway. The importance of this limitation depends on the characteristics of the compounds being tested. As was shown recently, the paracellular pathway seems very important for hydrophilic compounds such as metformin.[119] Since the paracellular pathway is important for certain classes of compounds, the decision to ignore this pathway could be risky. At this stage, important characteristics of a compound that may be relevant to its transepithelial permeation are not known.

PAMPA has several disadvantages such as lack of paracellular transport, enzymes, and membrane transporters. These disadvantages, together with the absence of cellular architecture makes it impossible to evaluate the role of discrete steps such as apical uptake, basolateral efflux, diffusion through the cytoplasm, and interaction with intracellular organelles, on transepithelial permeation. Given the important differences between PAMPA and Caco-2 cells the logical approach regarding the use of these systems is to consider them as complementary rather than as equivalent alternatives (Figure 9.4). The simplicity of PAMPA makes it a useful tool in early drug discovery when there is interest in assessing the permeability potential of large numbers of compounds. Caco-2 cells are more suitable during lead optimization or preclinical development. At these stages of the discovery/development process there is a greater need to determine true transepithelial permeability coefficients that can be used to gain insight into mechanistic aspects of drug absorption or to predict drug absorption in an intact organism. The correct placement of PAMPA and Caco-2 permeability measurements to capitalize on the natural complementary utility of these models was described recently.[120]

As is the case with any experimental model it is important to make sure that PAMPA is used for the right type of experiments. Considering the properties of the system it is clear that the value of PAMPA is an early discovery to obtain approximate permeability parameters that can indicate the potential of the compound to permeate cell membranes.

Recent studies have found a good correlation between intrinsic permeability coefficients in the PAMPA system and % absorption in humans for a selected series of compounds.[117] In trying to help scientists understand the potential utility of PAMPA, it is important to avoid making generalization of the predictive value of this technique based on data from small series of compounds.[117,118,120] In addition, considering the well-known interlaboratory variability in Caco-2 permeability values, and in the absence of standardized Caco-2 permeability values, it is difficult to make general conclusions regarding PAMPA–Caco-2 correlations, using Caco-2 data from one laboratory.

9.8 *In Silico* Methods to Predict Drug Absorption

Oral dosing is the preferred route of drug administration, but the labor involved in generating experimental data to estimate drug absorption potential is substantial. Therefore, the possibility of developing computer programs capable of predicting drug absorption is very attractive. The simplest approach consists of applying Lipinski's rule of 5.[121] This rule is not used to predict drug absorption but as a filter to flag compounds with potential absorption problems. Other computational strategies consist of predicting the absorption or permeability of molecules based only on molecular descriptors.[122,123] The progress of this strategy is very slow because permeability and absorption are not basic properties of the molecules determined by the chemical structure, and therefore predictable by using chemical descriptors alone. Permeability and absorption depend on variables, such as pH, formulation/excipients, and site of administration relative to site of absorption.[124] Despite this intrinsic difficulty researchers have predicted permeability coefficients (P_{app}) or fraction absorbed (F_{abs}) from the chemical structure of the compounds with varying degrees of success.[122–126] Molecular descriptors found to be useful in predicting P_{app} or F_{abs} are H bonds and dynamic polar surface area (PSAd).[127,128] Regardless of the approach taken to predict permeability and absorption, and considering the small size of the databases used to generate these models, it is clear that one of the main reasons for the slow progress in this area is the lack of sufficiently large databases of suitable quality for model development. In the absence of suitable data, it is extremely difficult to develop models to predict permeability or absorption outside the chemical space contained in the training set. Unless the models are derived from diverse and reliable databases, their utility will probably remain limited to compounds closely related to the training sets used to develop the models.

Although many laboratories have shown a correlation between their Caco-2 P_{app} data and F_{abs} in humans,[35–39] the size of most data sets is too small to be able to derive useful permeability models. Thus, the possibility of pooling data from different sources to increase the size of the database used for modeling seems reasonable; however, the success of this approach is highly unlikely considering the magnitude of the interlaboratory variability in P_{app} values.[40] The possibility of developing useful *in silico* models to predict absorption and permeability will remain limited unless databases of appropriate quality are developed.

9.9 Final Remarks

Caco-2 cells have played a major role in permeability studies during the last decade; this role has evolved substantially over time. Originally, the main advantages of Caco-2 cells were that they permitted greater versatility in experimental design, required smaller amount of material, experiments could be shorter and the sample matrices were simple, which simplified drug analysis. The relevance of Caco-2 cells as a model system of intestinal permeability was bolstered by their expression of numerous transporters and drug-metabolizing enzymes. As the cells became a key component of the arsenal of pharmaceutical scientists the systems used to cultured the cells underwent multiple and drastic modifications to increase throughput. These modifications, which included automation and miniaturization, together with LC-MS technology enhanced the

throughput of the Caco-2 cells permeability assay greatly. The Caco-2 cell system is largely responsible for the radical increase in awareness of the importance of permeability on the ultimate success of drug candidates among medicinal chemists and drug metabolism scientists.

With the advent of simpler, higher throughput techniques such as PAMPA focused on obtaining approximate permeability estimates for large number of compounds, the role of Caco-2 cells in the drug discovery stream continues to change. PAMPA is used in many settings as the frontline permeability screening tools for large numbers of compounds and Caco-2 cells is used for more detailed studies where the emphasis is to obtain mechanistic information or a P_{app} value that can be related to an *in vivo* parameter such as F_{abs}. In addition, Caco-2 cells are finding greater applications in the study of membrane transporters and for BCS drug classification. While the idea of predicting drug absorption and permeability constitutes a highly desirable goal, the incomplete understanding of the complex processes that control drug permeation and the lack of databases of suitable size and robustness will make this task difficult to achieve in the near future.

Abbreviations

HTS	high-throughput screening
CYP	cytochrome P_{450}
UDPGT	UDP glucuronyltransferase
PEPT1	peptide transporter 1
ACE	angiotensin-converting enzyme inhibitor
ABC	ATP-binding cassette
MDR	multidrug resistance
Pgp	P-glycoprotein
MRP	multidrug-resistance associated protein
cMOAT	canalicar multispecific organic anion transporter
Log D/P	logarithm of the octanol/water (or buffer) distribution/partition coefficient
MDCK	Madin–Darby canine kidney
BBMV	Brush–Border membrane vesicles
TEER	transepithelial electrical resistance
PAMPA	parallel artificial membrane permeation assay
LC/MS	liquid chromatography/mass spectrometry (single quadrupole)
LC/MS/MS	liquid chromatography/mass spectrometry (triple quadrupole)
HPLC	high-performance liquid chromatography
UV	ultraviolet
DMSO	dimethylsulfoxide
BSA	bovine serum albumin

References

1. Neutra, M. and Kraehenhuhl, J.P., Transepithelial transport of proteins by intestinal epithelial cells, in *Biological Barriers to Protein Delivery,* Audus, K.L. and Raub, T.J., Eds., Plenum Press, New York, 1993, pp. 107–129.

2. Pade, V. and Stavchansky, S., Estimation of the relative contribution of the transcellular and paracellular pathway to the transport of passively absorbed drugs in the Caco-2 cell culture model, *Pharm. Res*., 14, 1210, 1997.
3. Sakai, M., Noach, A.B., Blom-Roosemalen, M.C., de Boer, A.G., and Breimer, D.D., Absorption enhancement of hydrophilic compounds by Verapamil in Caco-2 cell monolayers, *Biochem. Pharmacol*., 48, 1199, 1994.
4. Ranaldi, G., Islam, K., and Sambuy, Y., D-cycloserine uses an active transport mechanism in the human intestinal cell line Caco-2, *Antimicrob. Agent Chemother.*, 38, 1239, 1994.
5. Smith, P.L. Wall, D.A., Gochoco, C.H., and Wilson, G., (D) Routes of delivery: case studies. (5) Oral absorption of peptides and proteins, *Adv. Drug Deliv. Rev.*, 8, 253, 1992.
6. Hu, M. and Amidon, G.L., Passive and carrier-mediated intestinal absorption components of captopril, *J. Pharm. Sci.*, 77, 1007, 1988.
7. Fisher, R.B., Active transport of salicylate by rat jejunum, *Q. J. Exp. Physiol.*, 66, 91, 1981.
8. Hidalgo, I.J., Ryan, F.M., Marks, G.J., and Smith P.L., pH-dependent transepithelial transport of cephalexin in rabbit intestinal mucosa, *Int. J. Pharmaceut.*, 98, 83, 1993.
9. Hidalgo, I.J. and Li, J., Carrier-mediated transport and efflux mechanisms in Caco-2 cells, *Adv. Drug Del. Rev.*, 22, 53, 1996.
10. Fernandes, P.B., Technological advances in high-throughput screening, *Curr. Opin. Chem. Biol.*, 2, 597, 1998.
11. Rodrigues, A.D., Preclinical drug metabolism in the age of high-throughput screening: an industrial perspective, *Pharm. Res.*, 14, 1504, 1997.
12. Tarbit, M.H. and Berman, J., High-throughput approaches for evaluating absorption, distribution, metabolism and excretion properties of lead compounds, *Curr. Opin. Chem. Biol.*, 2, 411, 1998.
13. Pinto, M., Robine-Leon, S., Appay, M.D., Kedinger, M., Triadou, N., Dussaulx, E., Lacroix, B., Simon-Assmann, P., Haffen, K., Fogh, J., and Zweibaum, A., Enterocytic-like differentiation and polarization of the human colon adenocarcinoma cell line Caco-2 in culture, *Biol. Cell*, 47, 323, 1983.
14. Grasset, E., Pinto, M., Dussaulx, E., Zwebaum, A., and Desjeux, J.-F., Epithelial properties of human colonic carcinoma cell line Caco-2: electrical parameters, *Am. J. Physiol.*, 247, C260, 1984.
15. Hidalgo, I.J., Raub, T.J., and Borchardt, R.T., Characterization of the human colon carcinoma cell line (Caco-2) as a model system for intestinal epithelial permeability, *Gastroenterology*, 96, 736, 1989.
16. Hunter, J., Jepson, M.A., Tsuruo, T., Simmons, N.L., and Hirst, B.H., Functional expression of P-glycoprotein in apical membranes of human intestinal Caco-2 cells, *J. Biol. Chem.*, 268, 14991, 1993.
17. Dix, C.J., Hassan, I.F., Obray, H.Y., Shah, R., and Wilson, G., The transport of vitamin B_{12} through polarized monolayers of Caco-2 cells, *Gastroenterology*, 98, 1272, 1990.
18. Takanaga, H., Tamai, I., and Tsuji, A., pH-dependent and carrier-mediated transport of salicylic acid across Caco-2 cells, *J. Pharm. Pharmacol.*, 46, 567, 1994.
19. Said, H.M., Redha, R., and Nylander, W., A carrier-mediated Na^+ gradient-dependent transport for biotin in brush-border membrane vesicles, *Am. J. Physiol.*, 253, G631, 1987.
20. Kool, M., de Haas, M., Scheffer, G.L., Scheper, R.J., van Eijk, M.J.T., Juijn, J.A., Baas, F., and Borst, P., Analysis of expression of cMOAT (MRP2), MRP3, MRP4, and MRP5, homologues of the multidrug resistance-associated protein gene (*MRP1*), in human cancer cell lines, *Cancer Res.*, 57, 3537, 1997.
21. Hidalgo, I.J. and Borchardt, R.T., Transport of bile acids in a human intestinal epithelial cell line, Caco-2, *Biochim. Biophys. Acta,* 1035, 97, 1990.
22. Hidalgo, I.J. and Borchardt, R.T., Transport of a large neutral amino acid (phenylalanine) in a human intestinal epithelial cell line: Caco-2, *Biochim. Biophys. Acta*, 1035, 25, 1990.
23. Dantzig, A.A. and Bergin, L., Uptake of the cephalosporin, cephalexin, by a dipeptide transport carrier in the human intestinal cell line, Caco-2, *Biochim. Biophys. Acta*, 1027, 211, 1990.
24. Ng, K.-Y. and Borchardt, R.T., Biotin transport in a human intestinal cell line (Caco-2). *Life Sci.*, 53, 1121, 1993.

25. Nicklin, P.L., Irwin, W.J., Hassan, I.F., Mackay, M., and Dixon, H.B.F., The transport of acidic amino acids and their analogues across monolayers of human intestinal absorptive (Caco-2) cells in vitro, *Biochim. Biophys. Acta,* 1995, 1269, 176–186.
26. Baranczyk-Kuzma, A., Garren, J.A., Hidalgo, I.J., and Borchardt, R.T., Substrate specificity and some properties of phenol sulfotransferase from human intestinal Caco-2 cells, *Life Sci.*, 49, 1197, 1991.
27. Carriere, V., Lesuffleur, T., Barbat, A., Rousset, M., Dussaulx, E., Costet, P., de Waziers, I., Beaune, P., and Zweibaum, A., Expression of cytochrome P-450 3A in HT29-MTX and Caco-2 clone TC7, *FEBS Lett.*, 355, 247, 1994.
28. Putnam, W.S., Ramanathan, S., Pan, L., Takahashi, L.H., and Benet, L.Z., Functional characterization of monocarboxylic acid, large neutral amino acid, bile acid and peptide transporters, and P-glycoprotein in MDCK and Caco-2 cells, *J. Pharm. Sci.*, 91, 2622, 2002.
29. Pfrunder, A., Gutmann, H., Beglinger, C., and Drewe, J., Gene expression of CYP3A4, ABC-transporters (MDR1 and MRP1-MRP5) and hPXR in three different human colon carcinoma cell lines, *J. Pharm. Pharmacol.*, 55, 59, 2003.
30. Gutmann, K., Fricker, G., Torok, M., Michael, S., Beglinger, C., and Drewe, J., Evidence for different ABC-Transporters in Caco-2 cells modulating drug uptake, *Pharm. Res.*, 16, 402, 1999.
31. Xia, C.Q., Liu, N., Yang, D., Miwa, G., and Gan, L.-S., Expression, localization, and functional characteristics of breast cancer resistance protein in Caco-2 cells, *Drug Metab. Dispos.*, 33, 637, 2005.
32. Hirohashi, T., Suzuki, H., Chu, X.-Y., Tamai, I., Tsuji, A., and Sugiyama, Y., Function and expression of multidrug resistance-associated protein family in human colon adenocarcinoma cells (Caco-2), *J. Pharmacol. Exp. Ther.*, 292, 265, 2000.
33. Walter E. and Kissel, T., Heterogeneity in the human intestinal cell line Caco-2 leads to differences in transepithelial transport, *Eur. J. Pharm. Sci.*, 3, 215, 1995.
34. Cogburn, J.N., Donovan, M.G., and Schasteen, C.S., A model of human small intestinal absorptive cells. 1, Transport barrier, *Pharm. Res.*, 8, 210, 1991.
35. Artursson, P. and Karlsson, J., Correlation between oral drug absorption in humans and apparent drug permeability coefficients in human intestinal epithelial (Caco-2) cells, *Biochem. Biophys. Res. Comm.*, 175, 880, 1991.
36. Chong, S., Dando, S.A., Soucek, K.M., and Morrison, R.A., In vitro permeability through Caco-2 cells is not quantitatively predictive of in vivo absorption for peptide-like drugs absorbed via the dipeptide transport system, *Pharm. Res.*, 13, 120, 1996.
37. Yee, S., In vitro permeability across Caco-2 cells (colonic) can predict in vivo (small intestinal) absorption in man- fact or myth, *Pharm. Res.*, 14, 763, 1997.
38. Gres, M.-C., Julian, B., Bourrie, M., Meunier, V., Roques, C., Berger, M., Boulenc, X., Berger, Y., and Fabre, G., Correlation between oral drug absorption in humans, and apparent drug permeability in TC-7 cells, a human epithelial cell line: comparison with the parental Caco-2 cell line, *Pharm. Res.*, 15, 726, 1998.
39. Yazdanian, M., Glynn, S.L., Wright, J.L., and Hawi, A., Correlating partitioning and Caco-2 cell permeability of structurally diverse small molecular weight compounds, *Pharm. Res.*, 15, 1490, 1998.
40. Hidalgo, I.J., Assessing the absorption of new pharmaceuticals, *Curr. Top. Med. Chem.*, 1, 385, 2001.
41. Volpe, A.D., Faustino, P.J., and Yu, L.X., Towards standardization of an in vitro method of drug absorption, *Pharmacopeial Forum*, 27, July-August, 2001.
42. Artursson, P., Epithelial transport of drugs in cell culture. I: a model for studying the passive diffusion of drugs over intestinal absorptive (Caco-2) cells, *J. Pharm. Sci.*, 79, 476, 1990.
43. Rothen-Rutishauser, B., Braun, A., Gunthert, M., and Wunderli-Allenspach, H., Formation of multilayers in the Caco-2 cell culture model: a confocal laser scanning microscopy study, *Pharm. Res.*, 17, 460, 2000.
44. Alsenz, J. and Haenel E., Development of a 7-day, 96-well Caco-2 permeability assay with high-throughput direct UV compound analysis, *Pharm. Res.*, 20, 1961, 2003.
45. Balimani, P., Patel, K., Marino, A., and Chong, S., Utility of 96 well Caco-2 cell system for increased throughput of P-gp screening in drug discovery, *Eur. J. Pharmaceut. Biopharm.*, 58, 99, 2004.

46. Kansy, M., Senner, F., and Gubernator, K., Physicochemical high throughput screening: parallel artificial membrane permeability assay in the description of passive absorption processes, *J. Med. Chem.*, 41, 1007, 1998.
47. Garberg, P., Eriksson, P., Schipper, N., and Sjostrom, B., Automated absorption assessment using Caco-2 cells cultured on both sides of polycarbonate membranes, *Pharm. Res.*, 16, 441, 1999.
48. Caldwell, G.W., Easlick, S.M., Gunnet, J., Masucci, J.A., and Demarest, K., In vitro permeability of eight beta-blockers through Caco-2 monolayers utilizing liquid chromatography/electrospray ionization mass spectroscopy, *J. Mass Spectrom.*, 33, 607, 1998.
49. Markowska, M., Oberle, R., Juzwin, S., Hsu, C.-P., Gryszkiewicz, M., and Streeter, J., Optimizing Caco-2 cell monolayers to increase throughput in drug intestinal absorption analysis, *J. Pharmacol. Toxicol. Meth.*, 46, 51, 2001.
50. Watt, A.P., Morrison, D., and Evans, D.C., Approaches to higher-throughput pharmacokinetics (HTPK) in drug discovery, *Drug Discov. Today*, 5, 17, 2000.
51. White, R.E. and Manitpisitkul, P., Pharmacokinetic theory of cassette dosing in drug discovery screening, *Drug Metab. Dispos.*, 29, 957, 2001.
52. Frick, L.W., Adkison, K.L., Wells-Knecht, K.J., Woolard, P., and Higton, D.M., Cassette dosing: rapid in vivo assessment of pharmacokinetics, *Pharmaceut. Sci. Technol. Today*, 1, 12, 1998.
53. Bravo, S., Nielsen, C.U., Amstrup, J., Frokjaer, S., and Brodin, B., In-depth evaluation of Gly-Sar transport parameters as a function of culture time in the Caco-2 cell model, *Eur. J. Pharm. Sci.*, 21, 77, 2004.
54. Chong, S., Dando, S.A., and Morrison, R.A., Evaluation of Biocoat intestinal epithelium dirrerentiation environment (3-day cultured Caco-2 cells) as an absorption screening model with improved productivity, *Pharm. Res.*, 14, 1835, 1997.
55. Liang, E., Chessic, K., and Yazdanian, M., Evaluation of an accelerated Caco-2 cell permeability model, *J. Pharm. Sci.*, 89, 336, 2000.
56. Lentz, K.A., Hayashi, J., Lucisano, L.J., and Polli, J.E., Development of a more rapid, reduced serum culture system for Caco-2 monolayers and application to the biopharmaceutics classification system, *Int. J. Pharmaceut.*, 200, 41–51, 2000.
57. Yamashita, S., Konishi, K., Yamazaki, Y., Taki, Y, Sakane, T., Sezaki, H., and Furuyama, Y., New and better protocols for a short-term Caco-2 cell culture system, *J. Pharm. Sci.*, 91, 669, 2002.
58. Guidance for industry, waiver of in vivo bioavailability and bioequivalence studies for immediate-release solid oral dosage forms based on a biopharmaceutics classification system, August 2000, CDER/FDA (http://www.fda.gov/cder/guidance/3618fn1.htm).
59. Yazdanian, M., Briggs, K., Jankovsky, C., and Hawi, A., The "high solubility" definition of the current FDA guidance on biopharmaceutics classification system may be too strict for acidic drugs, *Pharm. Res.*, 21, 293, 2004.
60. Polli, J.E., Yu, L.X., Cook, J.A., Amidon, G.L., Borchardt, R.T., Burnside, B.A., Burton, P.S., Chen, M.-L., Conner, D.P., Faustino, P.J., Hawi, A., Hussain, A.S., Joshi, H., Kwei, G., Lee, V.H.-L., Lesko, L.J., Lipper, R.A., Loper, A.E., Nerurkar, S.G., Polli, J.W., Sanvordeker, D.R., Taneja, R., Uppoor, R.S., Vattikonda, C.S., Wilding, I., and Zhang, G., Summary workshop report: biopharmaceutics classification system-implementation challenges and extension opportunities, *J. Pharm. Sci.*, 93, 1375, 2004.
61. Beauchamp, L.M., Orr, G.F., de Miranda, P., Burnette, T., and Krenitsky, T.A., Amino acid ester prodrugs of acyclovir, *Antivir. Chem. Chemother.*, 3, 57, 1992.
62. Ganapathy, M.E., Huang, W., Wang, H., Ganapathy, V., and Leibach, F.H., Valacyclovir: a substrate for the intestinal and renal peptide transporters PepT1 and PepT2, *Biochem. Biophys. Res. Commun.*, 246, 470, 1998.
63. Tamai, I., Nakahishi, T., Nakahara, H., Sai, Y., Ganapathy, V., Leibach, F.H., and Tsuji, A., Improvement of L-dopa absorption by dipeptidyl derivation, utilizing peptide transporter PepT1, *J. Pharm. Sci.*, 87, 1542, 1998.
64. Tomita, M., Sawada, T., Ogawa, T., Ouchi, H., Hayashi, M., and Awazu, S., Differences in the enhancing effects of sodium caprate on colonic and jejunal drug absorption, *Pharm. Res.*, 9, 648, 1992.

65. Sancho-Chust, V., Bengochea, M., Fabra-Campos, S., Casabo, V.G., Martinez-Camara, M.J., and Martin-Villodre, A., Experimental studies on the influence of surfactants on intestinal absorption of drugs. Cefadroxil as model drug and sodium lauryl sulfate as model surfactant: studies in rat duodenum, *Arzneim.-Forsch./Drug Res.*, 45, 1013, 1995.
66. Constantinides, P., Lancaster, C.M., Marcello, J., Chiossone, D.C., Orner, D., Hidalgo, I.J., Smith, P.L., Sarkahian, A.B., Yiv, S.H., and Owen, A.J., Enhanced intestinal absorption of an RGD peptide from water-in-oil microemulsions of different composition and particle size, *J. Control. Rel.*, 34, 109, 1995.
67. Raiman, J., Niemi, R., Vepsalainen, J., Yritys, K., Jarvinen, T., and Monkkonen, J., Effects of calcium and lipophilicity on transport of clodronate and its esters through Caco-2 cells, *Int. J. Pharm.*, 213, 135, 2001.
68. Rege, B.D., Yu, L.X., Hussain, A.S., and Polli, J.E., Effect of common excipients on Caco-2 transport of low-permeability drugs, *J. Pharm. Sci.*, 90, 1776, 2001.
69. Takahashi, Y., Kondo, H., Yasuda, T., Watanabe, T., Kobayashi, S.-I., and Yokohama, S., Common solubilizers to estimate the Caco-2 transport of poorly water-soluble drugs, *Int. J. Pharm.*, 246, 85, 2002.
70. McGann, L.E., Kinetics of osmotic water movement in chontrocites isolated from articular cartilage and application to cryopreservation, *J. Orthopedic Res.*, 6, 109, 1988.
71. Wolfinbarger, L. and Hu, J.-F., DMSO concentration in fresh cryopreserved porcine valved conduit tissue, *Cryobiology*, 31, 461, 1994.
72. Demirbas, S. and Stavchansky, S., Effects of citicholine and dimethylsulfoxide on transepithelial transport of passively diffused drugs in the Caco-2 cell culture model, *Int. J. Pharm.*, 251, 107, 2003.
73. Da Violante, G., Zerrouk, N., Richard, I., Provot, G., Chaumeil, J.C., and Arnaud, P., Evaluation of the cytotoxicity effect of dimethylsulfoxide (DMSO) on Caco2/TC7 colon tumor cell cultures, *Biol. Pharm. Bull.*, 25, 1600, 2002.
74. Putnam, W.S., Ramanathan, S., Pan, L., Takahashi, L., and Benet, L.Z., Functional characterization of monocarboxylic acid, large neutral amino acid, bile acid and peptide transporters, and P-glycoprotein in MDCK and Caco-2 cells, *J. Pharm. Sci.*, 91, 2622, 2002.
75. Lennernas, H., Does fluid flow across the intestinal mucosa affect quantitative oral drug absorption? Is it time for a reevaluation?, *Pharm. Res.*, 12, 1573, 1995.
76. Fine, K.D., Santa Ana, C.A., Porter, J.L., and Fordtran, J.S., Effect of D-glucose on intestinal permeability and its passive absorption in human small intestine in vivo, *Gastroenterology*, 105, 1117, 1993.
77. Soergel, K.H., Showdown at the tight junction, *Gastroenterology*, 105, 1247, 1993.
78. Knipp, G.T., Paracellular diffusion in Caco-2 cell monolayers: effect of perturbation on the transport of hydrophilic compounds that vary in charge and size, *J. Pharm. Sci.*, 86, 1105, 1997.
79. Artursson, P. and Magnusson, C., Epithelial transport of drugs in cell culture. II. Effect of extracellular calcium concentration on the paracellular transport of drugs of different lipophilicities across monolayers of intestinal epithelial (Caco-2) cells, *J. Pharm. Sci.*, 79, 595, 1990.
80. Boulenc, X., Roques, C., Joyeux, H., Berger, Y., and Fabre, G., Biophosphonates increase tight junction permeability in the human intestinal epithelial (Caco-2) model, *Int. J. Pharm.*, 123, 13, 1995.
81. Gan, L.-S., Hysu, P.-H., Pritchard, J.F., and Thakker, D., Mechanism of intestinal absorption of ranitidine and ondansetron: transport across Caco-2 cell monolayers, *Pharm. Res.*, 10, 1722, 1993.
82. Tanaka, Y., Taki, Y., Sakane, T., Nadai, T., Sezaki, H., and Yamashita, S., Characterization of drug transport through tight-junctional pathway in Caco-2 monolayer: comparison with isolated rat jejunum and colon, *Pharm. Res.*, 12, 523, 1995.
83. Pappenheimer, J.R. and Reiss, K.Z., Contribution of solvent drag through intercellular junctions to absorption of nutrients by the small intestine of the rat, *J. Membr. Biol.*, 100, 123, 1987.
84. Madara, J.L. and Pappenheimer, J.R., Structural basis for physiological regulation of paracellular pathways in intestinal epithelia, *J. Membr. Biol.*, 100, 149, 1987.
85. Tsuji, A. and Tamai, I., Carrier-mediated intestinal transport of drugs, *Pharm. Res.*, 13, 963 1996.

86. Bai, J.P.F., Hu, M., Subramanian, P., Mosberg, H.I., and Amidon, G.L., Utilization of peptide carrier system to improve intestinal absorption: targeting prolidase as a prodrug-converting enzyme, *J. Pharm. Sci.*, 81, 113, 1992.
87. Borchardt, R.T. and Hu, M., Transport of a large neutral amino acid in a human intestinal epithelial cell line (Caco-2): uptake and efflux of phenylalanine, *Biochim. Biophys. Acta,* 1135, 233, 1990.
88. Trotter, P.J. and Storch, J., Fatty acid uptake and metabolism in a human intestinal cell line (Caco-2): comparison of apical and basolateral incubation, *J. Lipid Res.*, 32, 293, 1991.
89. Ng, K.-Y. and Borchardt, R.T., Biotin transport in a human intestinal epithelial cell line (Caco-2), *Life Sci.*, 53, 1121, 1993.
90. Woodcock, S., Williamson, I., Hassan, I., and Mackay, M., Isolation and characterization of clones from the Caco-2 cell line displaying increased taurocholic acid transport, *J. Cell Sci.*, 98, 323, 1991.
91. Peterson, M. and Mooseker, M.S., Characterization of the enterocyte-like brush border cytoskeleton of the C2Bbe clones of the human intestinal cell line, Caco-2, *J. Cell Sci*., 102, 581, 1992.
92. Horie, K., Tang, F., and Borchardt, R.T., Isolation and characterization of Caco-2 subclones expressing high levels of multidrug resistance protein efflux transporter, *Pharm. Res.*, 20, 161, 2003.
93. Hu, M., Subramanian, P., Mosberg, H.I., and Amidon, G.L., Use of the peptide carrier system to improve the intestinal absorption of L-alpha-methyldopa: carrier kinetics, intestinal permeabilities, and in vitro hydrolysis of dipeptidyl derivatives of L-alpha-methyldopa, *Pharm. Res.*, 6, 66, 1989.
94. Han, H.-K., de Vrueh, R.L.A., Rhie, J.K., Covitz, K.-M.Y., Smith, P.L., Lee, C.-P., Oh, D.-M., Sadee, W., and Amidon, G.L., 5′-Amino acid esters of antiviral nucleoside, acyclovir, and AZT are absorbed by the intestinal PEPT1 peptide transport, *Pharm. Res*., 15, 1154, 1998.
95. Steingrimsdottir, H., Gruber, A., Palm, C., Grimfors, G., Kalin, M., and Eksborg, S., Bioavailability of acyclovir after oral administration of acyclovir and its prodrug valaciclovir to patients with leukopenia after chemotherapy, *Antimicrob. Agents Chemother.*, 44, 207, 2000.
96. Prueksaritanont, T., Gorham, L.M., Hochman, J.H., Tran, L.O., and Vyas, K.P., Comparative studies of drug-metabolizing enzymes in dog, monkey, and human small intestines, and in Caco-2 cells, *Drug Metab. Dispos.*, 24, 634, 1996.
97. Boulenc, X., Bourrie, M., Fabre, I., Roque, C., Joyeux, H., Berger, Y., and Fabre, G., Regulation of cytochrome P4501A1 gene expression in a human intestinal cell line, Caco-2, *J. Pharmacol. Exp. Ther.*, 263, 1471, 1992.
98. Raeissi, S.D., Guo, Z., Dobson, G.L., Artursson, P., and Hidalgo, I.J., Comparison of CYP3A activities in a subclone of Caco-2 cells (TC7) and human intestine, *Pharm. Res.* 14, 1019, 1997.
99. Yashioka, M., Erickson, R.H., Matsumoto, H., Gum, E., and Kim, Y.S., Expression of dipeptidyl aminopeptidase IV during enterocytic differentiation of human colon cancer (Caco-2) cells, *Int. J. Cancer*, 47, 916, 1991.
100. Munzel, P.A., Schmohl, S., Heel, H., Kalberer, B., Bock-Hennig, B.S., and Bock, K.W., Induction of human UDP glucuronosyltransferases (UGT1A6, UGT1A9, and UGT2B7) by *t*-butylhydroquinone and 2,3,7,8-tetrachlorodibenzo-*p*-dioxin in Caco-2 cells, *Drug Metab. Dispos.*, 27, 569, 1999.
101. Schmiedlin-Ren, P., Thummel, K.E., Fisher, J.M., Paine, M.F., Lown, K.S., and Watkins, P., Expression of enzymatically active CYP3A4 by Caco-2 cells grown on extracellular matrix-coated permeable supports in the presence of 1α,25-dihydrovitamin D_3, *J. Pharmacol. Exp. Ther*., 51, 741, 1997.
102. Crespi, C.L., Penman, B.W., and Hu, M., Development of Caco-2 cells expressing high levels of cDNA-derived cytochrome P4503A4, *Pharm. Res.*, 13, 1635, 1996.
103. Ranaldi, G., Islam, K., and Sambuy, Y., Epithelial cells in culture as a model for the intestinal transport of antimicrobial agents, *Antimicrob. Agent Chemother.*, 36, 1374, 1992.
104. Irvine, J.D., Takahashi, L., Lockhart, K., Cheong, J., Tolan, J.W., Selick, H.E., and Grove, J.R., MDCK (Madin-Darby canine kidney) cells: a tool for membrane permeability screening, *J. Pharm. Sci.*, 88, 28, 1999.

105. Wacher, V.J., Wu, C.-Y., and Benet, L.Z., Overlapping substrate specificities and tissue distrivution of cytochrome P450 3A and P-glycoprotein: implications for drug delivery and activity in cancer chemotherapy, *Mol. Carcinog.*, 13, 129, 1995.
106. Tran, C.D.H., Timmins, P., Conway, B.R., and Irwin, W.J., Investigation of the coordinated functional activities of cytochrome P450 3A4 and P-glycoprotein in limiting the absorption of xenobiotics in Caco-2 cells, *J. Pharm. Sci.*, 91, 117, 2002.
107. Cummins, C.L., Jacobsen, W., and Benet, L.Z., Unmasking the dynamic interplay between intestinal P-glycoprotein and CYP3A4, *J. Pharmacol. Exp. Ther.*, 300, 1036, 2002.
108. Hochman, J.H., Chiba, M., Yamazaki, M., Tang, C., and Lin, J.H., P-glycoprotein-mediated efflux of indinavir metabolites in Caco-2 cells expressing cytochrome P450 3A4, *J. Pharmacol. Exp. Ther.*, 298, 323, 2001.
109. Taipalensuu, J., Tornblom, H., Lindberg, G., Einarsson, C., Sjoqvist, F., Melhus, H., Garberg, P., Sjostrom, B., Lundgren, B., and Artursson, P., Correlation of gene expression of ten drug efflux proteins of ATP-binding cassette transport family in normal human jejunum and in human intestinal epithelial Caco-2 cell monolayers, *J. Pharmacol. Exp. Ther*., 299, 164–170, 2001.
110. Cummins, C.L., Salphati, L., Reid, M.J., and Benet, L.Z., In vivo modulation of intestinal CYP3A metabolism by P-glycoprotein: studies using the rat single-pass intestinal perfusion model, *J. Pharmacol. Exp. Ther.*, 305, 306, 2003.
111. Tamura, S., Tokunaga, Y., Ibuki, R., Amidon, G.L., Sezaki, H., and Yamashita, S., The site-specific transport and metabolism of tacrolimus in rat small intestine, *J. Pharmacol. Exp. Ther.*, 306, 310, 2003.
112. Quan, Y., Hattori, K., Lundborg, E., Fujita, T., Murakami, M., Muranishi, S., and Yamamoto, A., Effectiveness and toxicity screening of various absorption enhancers using Caco-2 cell monolayers., *Biol. Pharm. Bull.*, 21, 620, 1998.
113. Stewart, B.H., Chan, O.H., Lu, R.H., Reyner, E.L., Schmid, H.L., Hamilton, H.W., Steinbaugh, B.A., and Taylor, M.D., Comparison of intestinal permeabilities determined in multiple in vitro and in situ models: relationship to absorption in humans, *Pharm. Res*., 12, 693, 1995.
114. Lennernas, H., Palm, K., Fagerholm, U., and Artursson, P., Comparison between active and passive drug transport in human intestinal epithelial (Caco-2) cells in vitro and human jejunum in vivo, *Int. J. Pharm*., 127, 103, 1996.
115. Dias, V. and Yatscoff, R.W., An in vitro method for predicting in vivo oral bioavailability of novel immunosuppressive drugs, *Clin. Biochem*., 29, 43, 1996.
116. Lau, Y.Y., Chen, Y.-H., Liu, T.-T., Li, C., Cui, X., White, R.E., and Cheng, K.-C., Evaluation of a novel in vitro Caco-2 hepatocyte hybrid system for predicting in vivo oral bioavailability, *Drug Metab. Dispos*., 32, 937, 2004.
117. Kansy, M., Senner, F., and Gubernator, K., Physicochemical high throughput screening: parallel artificial permeation assay in the description of passive absorption processes, *J. Med. Chem*., 41, 1007, 1998.
118. Bermejo, M., Avdeef, A., Ruiz, A., Nalda, R., Ruell, J.A., Tsinman, O., Gonzalez, I., Fernandez, C., Sanchez, G., Garrigues, T.M., and Merino, V., PAMPA — a drug absorption in vitro model 7. Comparing rat in situ, Caco-2, and PAMPA permeability of fluoroquinolones, *Eur. J. Pharm. Sci*., 21, 429, 2004.
119. Saitoh, R., Sugano, K., Takata, N., Tachibana, T., Higashida, A., Nabuchi, Y., and Aso, Y., Correction of permeability with pore radius of tight junctions in Caco-2 monolayers improves the prediction of the dose fraction of hydrophilic drugs absorbed by humans, *Pharm. Res*., 21, 749, 2004.
120. Kerns, E.H., Li, D., Petusky, S., Farris, M., Ley, R., and Jupp, P., Combined application of parallel artificial membrane permeability assay and Caco-2 permeability assays in drug discovery, *J. Pharm. Sci*., 93, 1440, 2004.
121. Lipinski, C.A., Lombardo, F., Dominy, B.W., and Feeney, P.J., Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, *Adv. Drug Del. Rev.*, 23, 3, 1997.
122. Palm, K., Luthman, K., Ungell, A.-L., Strandlund, G., and Artursson, P., Correlation of drug absorption with molecular surface properties, *J. Pharm. Sci.*, 85, 32, 1996.

123. Klopman, G., Stefan, L.R., and Saiakhov, R.D., ADME Evaluation 2. A computer model for the prediction of intestinal absorption in humans, *Eur. J. Pharm. Sci.*, 17, 253, 2002.
124. Burton, P.S., Goodwin, J.T., Vidmar, T.J., and Amore, B.M., Predicting drug absorption: how nature made it a difficult problem, *J. Pharmacol. Exp. Ther.*, 303, 889, 2002.
125. Parrott, N. and Lave, T., Prediction of intestinal absorption: comparative assessment of GASTROPLUS™ and IDEA™, *Eur. J. Pharm. Sci.*, 17, 51, 2002.
126. Stenberg, P., Norinder, U., Luthman, K., and Artursson, P., Experimental and computational screening models for the prediction of intestinal drug absorption, *J. Med. Chem.*, 44, 1927, 2001.
127. Palm, K., Stenberg, P., Luthman, K., and Artursson, P., Polar molecular surface properties predict the intestinal absorption of drugs in humans, *Pharm. Res.*, 14, 568, 1997.
128. Oprea, T.I. and Gottfries, J., Toward minimalistic modeling of oral drug absorption, *J. Mol. Graphics Mod.*, 17, 261, 1999.

10

The Promise of Metabonomics in Drug Discovery

Harold J. Kwalwasser and Pauline Gee

CONTENTS

10.1 Introduction

The disciplines evolving from the sequencing of the human genome, transcriptomics and proteomics, have the potential to facilitate the drug discovery process but the greatest value lies in the integration of these areas of research focus. At the same time, another "omics" discipline has been significantly improving our decades-long knowledge of metabolites. Rather than quantifying one or two metabolites at a time, researchers involved in metabonomics (see Box 10.1) are attempting the comprehensive and quantitative study of changes in the metabolome (the molecules representative of the substrates, products, and intermediates of biochemical pathways on which proteins act). Historically, the focus of researchers has been a single gene or protein or metabolite. With the advances in the array concept, multiplexing of gene or protein expressions and support vector machine analyses of data generated from these platforms, clustering of such data has led to more predictive and robust interpretations. It is no surprise that interpretation of quantitative patterns of metabolites should follow since small molecules can be monitored in bodily fluids obtained with minimal or no invasion of the body. Traditionally, clinicians have measured single-metabolite markers and have intuitively bunched several together for confirmation or for eliminating a certain disease or metabolic condition. However, the greatest potential lies in the challenge of integrating all of these "omics" into a Systems Biology approach. With such an understanding, there is a promise of great value to drug discovery and development.

Box 10.1

Although the terms "metabonomics" and "metabolomics" are not always used consistently, Jeremy K. Nicholson of Imperial College and his colleagues coined the term "metabonomics"[1] in 1996 to refer to the "quantitative measurement of the dynamic multiparametric metabolic response of living systems to pathophysiological stimuli or genetic modification." According to this group, metabonomics is derived from "meta" and "nomics" derived from the Greek words for change and rules, respectively.[2] It does not measure the metabolism of the compound itself, but rather looks for a pattern of change in hundreds or thousands of endogenous biomolecules that is a "fingerprint" of efficacy or toxicity. "Metabolomics," is then reserved for the metabolic regulation and fluxes in individual cells or cell types and the measurement thereof. Others use the two terms interchangeably and as with any new field, nomenclature discussions are unavoidable, until we have a clear understanding of the boundaries. While this chapter attempts to follow the spirit of this distinction, the literature is not always in agreement.

The National Institutes of Health (NIH) has characterized "[t]he general aim of metabolomics … to identify, measure and interpret the complex, time-related concentration, activity and flux of endogenous metabolites in cells, tissues and other biosamples such as blood, urine and saliva."[3] Taken together, the three "omics" disciplines represent a systemic and powerful approach to mapping cellular networks, allowing for faster and more predictive biology in terms of risk assessment, potential therapeutic benefit, as well as diagnosis.

Metabonomics data complement gene expression and proteomic data. Whereas the latter reflect calculations of the potential for disease or toxicity, metabonomics reflects actual biological events. Focusing on metabolites allows monitoring a dynamic profile rather than a snapshot, creating a mechanistic level of understanding of toxicology.[4] Unlike transcriptomics and proteomics, metabonomics can give information on a whole organism's functional integrity over time after drug exposure.[5]

Moreover, the metabolome is more sensitive to changes in individual enzyme or transcript levels than the transcriptome or proteome. While metabolic fluxes are not always affected by changes in quantities of individual enzymes, when concentrations of individual metabolites do change, they are more closely linked than the transcriptome or proteome to variations in phenotype that are most relevant to human health. Often enzyme concentrations change to maintain concentrations of certain metabolites essential to physiological equilibrium associated with human health. Adelbert Roscher of the University of Munich succinctly characterized the relationship that metabolite profiling "measures the real outcome of potential changes suggested by genomics and proteomics."[6]

Metabonomics shows promise for other applications in drug discovery and development. Clinicians already use metabolic biomarkers to diagnose disease; it is not difficult to transfer the concept of patterns in gene expression to patterns of metabolites to describe and define the subtleties of disease that may be unique to a single individual. Furthermore, since stages of a disease may be further defined, treatments can be more rational and optimized. Genetic and proteinaceous changes could be compensated for elsewhere in the organism so that they are not necessarily part of the pathology. Many pathway feedback mechanisms are not reflected in protein concentration or gene expression changes. Mammalian metabolic control functions are dispersed across many cell types in topographically distinct locations that are in different physiological stages, raising serious questions about the value of simple pathway modeling. Environment and diet can also affect disease and toxicity, and both have an impact on the organism independent

of gene expression or proteomic activity. In addition, there must be an accounting for the entire gut microflora to understand the dynamics of drug efficacy and adverse reactions.[7]

10.2 Applications of Metabonomic Data

Metabonomic data have the promise of making significant contributions throughout the drug discovery and development process. During the discovery phase, metabonomic data can facilitate early *in vivo* toxicological testing, lead compound selection, prelead prioritization, and *in vivo* efficacy screening in animal models. During development, the data can find new preclinical and clinical safety biomarkers and mechanisms, and allow for metabotyping and the validation of animal models against human disease profiles. Furthermore, the adoption of this concept may be faster because the community has been measuring metabolites in bodily fluids for quite some time.

While not yet validated, Nicholson et al.,[5] believe that metabonomic data may be highly relevant in five key areas of drug discovery: (1) identifying the target mechanism; (2) deriving the biochemical mechanism of toxicity; (3) determining the combination of biomarkers for the onset, progression, and regression of a lesion; (4) providing metabolic fingerprints that can be used in the designs of trial and evaluation of animals as disease models; and (5) deriving new biochemically based assays for disease diagnosis and the identification of combination of biomarkers for disease.

Schwetz[8] foresees that metabonomic data will lead to a discussion of the distinction between physiologic adaptation and toxicity on the basis of new measures of proteomic and metabonomic changes. Such data improve the understanding of the relationship between changes in the physiologic metabolic capabilities of cells, tissues, and the whole body as it relates to responses to agents of exposure. Thus, improved criteria for measuring the fidelity of disease models should result in better disease models that will lead to more efficient drug discovery and development.

10.3 Acquisition of Metabolic Data

There is not only a controversy over nomenclature but there is also a comparable discussion over the number of metabolites in the metabolome. Almost by definition, any substance involved in metabolism either as a product of metabolism or necessary for metabolism should be considered, and estimates of 3000 major metabolites have been reported.[9] Others have indicated that any metabolite smaller than 1000 Da, which would incorporate virtually all energy pathways, all catabolic pathways, and many biosynthetic pathways, might be reasonable.[5]

It is this diversity of metabolites that presents the most difficult challenge. There are not only a variety of metabolites but they are also dynamic spatially and temporally. To master a universal "omics" approach, one has to have several measurement strategies. For example, a suite of quantitative methods analyzing key metabolites from different biochemical pathways and metabolic "fingerprinting" that compares patterns like multidimensional hierarchical clustering or principal cluster analysis.[9]

However, chemical diversity of metabolites requires a variety of detection platforms and presents a unique challenge when one analyzes the metabolome for fingerprinting or

clustering. Thus, each cluster or signature may be specific to a platform rather than to a physiologic state. No single platform, nuclear magnetic resonance (NMR), chromatography, or mass spectrometry (MS) can measure the universe of metabolites with sufficient specificity and sensitivity to produce large, balanced datasets to determine the rules that govern the changes during normal healthy physiologic states and those of disease or toxic states. To address these challenges, the NIH Roadmap Initiatives focus on new technologies or new strategies to use established technologies.[3] Thus, any discussion of measurement must focus both on improvements to individual measurement techniques and on strategies for their joint use in various combinations.

10.3.1 Mass Spectrometry and Chromatography

Mass spectrometry and gas or liquid chromatography (GC–LC) encompass a variety of methods used to isolate and measure metabolites. MS is good for broad screening because of its high sensitivity, high mass range, and high dynamic range. It is capable of identifying hundreds of individual molecular species in a single analysis. GC, which is most useful for metabolites below 800 kDa, while high-performance liquid chromatography (HPLC), yield rigorous quantification and can be as sensitive as MS. However, they cannot match MS' scope of measurements in a single assay, and they can be too slow for high-throughput analysis. Therefore, the methods are often used in combination, particularly to achieve the system goals of metabonomics.

Since 1980s, these various methods have evolved and have been refined so that they now provide highly sensitive measures for a range of metabolites. Because the goal of metabonomics is to measure as many metabolites as possible in an organism, various tools are often used in combination for the maximum coverage.

Recent developments in GC–LC–MS have made the following major improvements in measurement:[10]

- Gas chromatography time-of-flight mass spectrometry (GC-TOF) allows better deconvolution than do most GC–MS instruments because they can record spectra, and thus sample various methods quickly.
- Fourier-transform ion cyclotron resonance mass spectrometry (FT-ICR) is a very high-resolution mass spectral method (10^5 to 10^6, with mass accuracy better than 1 ppm), which allows separation and empirical formula measurement of potentially thousands of metabolites without chromatic separation.
- LC–MS has lower chromatographic resolution than GC–MS methods, but can access greater mass ranges because volatilization for the chromatographic step is unnecessary.
- Hydrophilic interaction liquid chromatography (HILIC) can handle the matrix effects nonvolatiles, which may reduce the evaporation of volatile ions during the electrospray process. Ion-exchange methods that require the use of salts could otherwise interfere with MS.
- Low-resolution electrospray MS allows for data obtainable in less than 1 min, allowing for enabling high-throughput, high-information screening.
- Matrix-assisted laser desorption ionization mass spectrometry (MALDI) uses a matrix, typically of aromatic acids, whose molecular weight is similar to those of typical metabolites and thus disallows the mass spectrometric measurement of the latter.

The application of more than one measurement technique, or multiple applications of one technique, at a time can materially enhance results. For example, GC–MS has been

routinely used for years. GC–MS is useful for the investigation of genetic metabolic diseases (through the analysis of organic acids and acylcarnitines), and has been successfully applied in the detection of hypocholesterolemia and Smith–Lemli–Optiz syndrome.[11] Another measurement strategy, the two-dimensional gas chromatography (GC × GC), will achieve orthogonal resolution for all separated peak clusters, rather than just a few selected ones. It has been used for drug analysis in doping control with 10 to 100 $\mu g/L^{-1}$ detection limits.[11] Like GC × GC, multidimensional HPLC (HPLC × HPLC) transfers specific peaks to the second dimension. The application of comprehensive HPLC × HPLC to small-molecule analyses has progressed since the demonstration of a separation of alcohol ethoxylates on aminopropyl silica columns, coupled to reverse-phase C_{18} columns.[12] Improvements in these systems are guided by optimization of the parameters that define the degree of correlation between separations in different dimension.[13] The range of measurement combinations is extensive. The various methods have now identified over 1000 metabolites.[14]

These methods have challenges and limitations. Sample preparation is required and can be elaborate, and dividing samples to analyze different classes of metabolites requires careful accounting. As noted above, GC and HPLC provide rigorous quantitative analyses, but they are often slow and ill-suited for high-throughput platforms. HPLC does not work well on nonpolar metabolites that do not ionize well or on very polar metabolite that do not separate well. Using GC for polar, nonvolatile metabolites requires less polar, volatile derivatives before they can be separated on a GC column. This issue of samples and sample preparation challenges the reproducibility of the measurements and emphasizes the importance of accurate metadata.

Particular measurement tools face further problems. FT-ICR will have to overcome ion–ion interferences at high concentrations. As noted above, LC–MS can be subject to matrix effects, such as the presence of nonvolatiles.

10.3.2 Nuclear Magnetic Resonance

Nuclear magnetic resonance was first used in the 1970s to study microbial metabolism.[15] When atoms of a molecule are placed in a strong magnetic field certain atoms absorb and then emit radio waves according to their natural resonant frequencies. Such resonant frequencies parallel the chemical environment where each element resides. Furthermore, the data from standard magnetic resonance imaging (MRI) instruments already in use can be adapted to spectroscopy for metabonomics by different acquisition software and a radiowave coil that can be placed in the body. Much of this work has been driven by the change in the pattern of metabolites of tumors responsive to a given treatment compared to those tumors that remain nonresponsive. This change can be measured long before anatomical changes.[16]

In comparison with other measurement methods, NMR has important strengths. Its great virtue is its noninvasive nature, "allowing one to obtain spatially resolved metabolic profiles and to investigate metabolomics *in vivo*."[10] There is little or no sample preparation. It is nondestructive. It is information-rich with regard to the determination of molecular structures because it can detect different chemical groups of metabolites simultaneously.

Phosphorus nuclei have been used for many years in *in vivo* NMR, especially for intracellular pH measurements. However, because most organic phosphates have similar chemical shifts, compound identification can be difficult without special attention being paid to culture conditions in the NMR tube.[15] Carbon NMR also yields significant results because of the large chemical shift dispersion and narrow lines of this nucleus. ^{13}C spectra reflect most of the chemical rearrangements that may take place between substrate and final product.

^{1}H-NMR is utilized extensively because protons are present in almost all metabolites in body fluids although urine allows better resolution than plasma because plasma lipids and proteins broaden the peaks in the low-frequency region of the spectrum. A ^{1}H-NMR spectrum provides a characteristic fingerprint of almost all proton-containing metabolites. The work in metabonomics has almost exclusively used ^{1}H-NMR for analyses of biofluids.[4] Toxicological studies have utilized ^{1}H-NMR of biofluids, such as urine where the toxins can produce changes in the concentration and pattern of metabolites that can provide information about the sites of the toxicity, the basic mechanism of toxicity, and potential biomarkers of organ-specific toxicity. A study by Nicholson[17] revealed different metabolite patterns in the 600 MHz ^{1}H-NMR spectra of rat urine following exposure to various toxins. For example, an increase in urinary glucose, amino acids, and various organic acids was indicative of damage to the proximal tubules in the renal cortex.

The most extensive study involving the use of ^{1}H-NMR has been undertaken by the Consortium for Metabonomic Toxicology (COMET), involving the Imperial College of Science, Technology and Medicine in London and six pharmaceutical companies.[18] The project focused on the development of comprehensive metabonomic databases and multivariate statistical models for prediction of toxicity, initially for liver and kidney toxicity in the rat and mouse. The study further assessed the analytical and biological variation that might arise through the use of metabonomics and the intersite reproducibility of the results, which was deemed robust.

Nicholson and Wilson[7] have also used ^{1}H-NMR measurement, together with discriminant partial least-squares analysis to distinguish various forms of coronary heart disease via blood samples, even though only a small number of samples was used. Consistent with the long tradition of studying metabolic pathways, Brecker et al.[19] monitored cosubstrates glucose and citrate and six end products during growth of *Escherichia coli* and Raamsdonk et al.[20] analyzed the metabolome in *Saccharomyces cerevisiae* to identify gene function. The central nervous system is the most challenging organ to develop and treat, and remains an important therapeutic area where the median age of populations increase.

Other nuclei, such as ^{13}C or ^{31}P, may be used to study other metabolite pools, or they can complement ^{1}H-NMR to create more sophisticated NMR spectra. ^{13}C-NMR provides a greater spectral range (~200 ppm) than ^{1}H-NMR (~15 ppm). Although lower natural abundance of ^{13}C (1.1%) yields lower sensitivity, it also provides an opportunity to use isotopic enrichment to trace specific metabolic pathways with enhanced sensitivity.[4] ^{31}P can observe high-energy phosphate metabolites such as adenosine triphosphate.

Techniques that observe heteronuclear correlations, such as ^{1}H–^{13}C, have significant possibilities. Polarization transfer techniques combined with inverse detection methods provide noticeable enhancements in NMR sensitivity for nuclides such as ^{13}C. Such multidimensional NMR can be useful in studying the biodegradation of xenobiotics because the pathways are unknown and the final and intermediate metabolites have to be established.[15] Using high-resolution magic-angle spinning (HRMAS) ^{1}H- and ^{13}C-NMR spectra were acquired in brain slices to demonstrate that compartmentation of metabolites could be resolved according to cell type as well as relative concentrations within each cell type.[21]

Other problems have been improved by various technological enhancements to ^{1}H-NMR itself. The usefulness of ^{1}H-NMR has suffered from a small spectral dispersion of the nucleus and an intense water signal originating from cell suspension or extracts. The situation is getting better as the machines migrate toward using stronger magnetic fields of 11.7 T (500 MHz proton frequency) or 14.1 T (600 MHz). ^{1}H-NMR at 500 to 600 MHz can detect small metabolite concentrations of less than 100 μM in 7 to 11 min, but at 800 to 900 MHz, it could reduce acquisition times sufficiently to do 300 samples per day.[4] The use of stronger field gradients will compensate for the fact that the longer the spin, the more

likely the diffusion. Using a strong field gradient >12 T/m^{-1}, various researchers report has yielded spatial resolution of a few micrometers.[10]

There have also been recent improvements to reduce background noise, which is caused by random thermal motions of electrons in the detection coil and in the first amplifier range.[15] Cryoprobes in which the coil and first amplifier are cooled to ~30 K and the size of sample coil reduced lower electrical resistance and noise. A further benefit noted by Grivet et al.[15] is that: "the smaller the probe, the closer the conductors are to the sample and the higher the measure of the coil's [efficiency], as 'felt' by the nuclei."

Also, adapting the coils depending on whether the experiment is concentration limited, or mass limited, can improve measurements. For example, when measuring the ATP content of living cells, viable cell suspension comprises at most ~10^{11} cells/mL and the ATP concentration is rarely >20 m*M*. Therefore, a larger sample means a strong signal, suggesting tube diameters of 20 mm. For mass-limited experiments, capillary sample tubes may be 1.7 mm. However, rather than using typical saddle coil probes in such a case, researchers have obtained much larger signal to noise ratios with microcoils directly wound on the sample tube.[15] Lacey et al.[22] report reducing a normalized limit of detection of 3 for proton spectroscopy at 600 MHz, single scan, to 0.54.

NMR may be limited by the width of the lines one wishes to observe. The magnetic inhomogeneity of the material surrounding a liquid sample, or the inhomogeneity of the sample itself, may cause broadening. Magic angle spinning for high resolution (HR-MAS), in which the sample tube is spun about its axis at an angle of 54.7° to the magnetic field direction at a rate several time 1000 Hz, can improve the measurement. MAS removes the line broadening caused by dipolar couplings, chemical shift anisotropies, and anisotropic magnetic susceptibility artifacts. Tissue metabolites already enjoy a degree of molecular mobility such that the line broadenings are greatly reduced from their true solid values; this means that only modest rotation rates (<10 kHz) are required. It was used heavily in the COMET project.[18]

10.3.3 Combining Nuclear Magnetic Resonance and Mass Spectrometry and Chromatography

Nuclear magnetic resonance traditionally has had low sensitivity and spectral resolution. It can provide rigorous quantification of analytes, but not accurate qualitative identification. Individual resonances may be sensitive to the chemical and physical environment of the molecule, which then requires appropriate preparation of samples. The process also has little dynamic range, in contrast to GC–MS.

As a consequence, using NMR simultaneously with MS can provide complementary data and can combine the power of LC–MS and LC–NMR to identify components specifically. Stopped flow achieves better detection while parallel arrangements enables either the NMR or the MS to be used as the intelligent detector in finding eluting peaks of interest.[23] As an example of the current combined approach, a computer-controlled post-LC-column splitter may create the NMR–MS interface. It directs 90 to 95% of the flow to the NMR and provides appropriate dilution and makeup flow for optimal MS ionization. The system has been enhanced by a loop storage mode, where portions of the sample can be stored in a delay loop before transfer to NMR. Chromatographic peaks can be trapped on line, in multiple-loop cassettes, stored, and then automatically transferred to for later NMR–MS analyses under computer automation.[23] The system provides a versatile analytical platform for complex mixture analysis and can benefit the pharmaceutical industry by facilitating drug metabolite identification. In the future, development of various combination systems, such as MS-directed LC–SPE–NMR–MS systems (SPE, solid-phase extraction)

equipped with cryogenic cooling of the NMR radio frequency coils, should provide a powerful analytical tool for drug discovery.[11]

10.4 Analyzing the Data

Metabonomics has changed the way data are analyzed. Instead of the single biomarker concept, the goal is to simultaneously measure the entire range of plasma metabolites to see changes in the patterns of the biomarkers. This requires both analytical platforms that can assess the broad range of metabolites in an individual and the knowledge of metabolism that is necessary to interpret the analytical data.

Initially, raw data must be deconvoluted. Kell[10] has identified three possibilities: (1) working out the signal from metabolites that are imperfectly separated using a hyphenated chromatographic method such as GC–MS, and hence their concentration; (2) providing a chemical identity for metabolites reproducibly recognized as being present as judged by, for example, their retention index and mass spectrum; and (3) using the metabolomic data to reconstruct the metabolic networks in which they participate.

There has been considerable work recently in the first area. New publicly available algorithms can process raw hyphenated data and produce lists of metabolites as outputs. For example, MSFACTs (metabolomics spectral formatting, alignment, and conversion tools) is a custom software package for alignment of GC–MS and LC–MS data to allow visualization using commercial statistical packages.[24] A high-speed peak alignment algorithm was developed and was shown to increase the efficacy of pattern recognition methods that is essential if the data are to used to make decisions in drug discovery and development.[25]

Unfortunately, progress in the second area, chemical identities for metabolites, has been more challenging. While there has been some considerable progress on identifying organic chemical metabolites, much more work will be needed to catalog and construct a database of all natural biological metabolites, both endogenous and those from the microflora.

There have been significant developments in reconstructing metabolic networks. Much of this work deals with pattern recognition strategies that are either unsupervised or supervised learning techniques. In the unsupervised technique, the system is shown as a set of inputs with no prior knowledge of the system assumed and then left to cluster metabolite data. In supervised learning, the desired responses associated with each of the inputs are known and the goal is to find a mathematical transformation (model) that will correctly associate all or some of the inputs with the target traits.[26] The basic idea with supervised learning is that there are some patterns (metabolic fingerprints) that have known desired responses. Nicholson[17] has reported on a comparison of the results of using a principal component analysis (PCA) unsupervised model with a supervised approach using soft independent modeling of class analogy (SIMCA) for a series of liver and renal toxins. The data illustrate how a supervised model can be used to differentiate hyrrazine type toxicity from renal toxins.

Reconstructing network organization is a challenge because many organisms will metabolize xenobiotics, opening up the possibility of as many as 200,000 metabolites in the plant kingdom.[10] However, annotated genomic data, such as that which can be found in the Kyoto Encyclopedia of Genes and Genomes (KEGG), can provide the baseline of reactions, although it would be helpful to have organism-specific ones. Work on *E. coli* and *S. cerevisiae* have model systems for some of these data because the number of metabolites is in the hundreds, rather than the thousands.

Some factors do make the process easier. Several studies suggest that metabolic networks are scale-free. There are also important stoichiometric constraints and elementary nodes

that restrict both the networks and the regulations that are possible. Various strategies to facilitate reconstruction have been reviewed by Kell[10] One strategy combines the networks that are reconstructed qualitatively from the genomic data with the constraints imposed by quantitative mass balancing. Subject to constraints about how cells choose or have evolved to regulate their fluxes in terms of growth rate or growth yield, it is possible to make some extremely powerful and successful predictions of whole-cell behavior at the metabolic and physiological levels from such *in silico* analysis alone. Mangan and Alon[27] developed network motifs that are arrangements of reactions, which include feedback structures that occur in metabolic pathways and, therefore, are assumed to have functional use. A development of metabolic control analysis (MCA), called coresponse analysis, has been useful in pattern recognition analysis of the metabolome and in recognizing that the covariation of pairs under different conditions can provide information of their connectedness. For a review of how to relate metabolite covariance matrices, using statistical and machine learning methods in deconstructing the highly multivariate data that metabolomics provides see the work of Fiehn and Weckwerth[28] from the Fiehn Laboratory. Finally, actively perturbing metabolic networks is a further strategy, but is obviously dependent on deciding which perturbation to make. Vance et al.[29] presented a computational strategy for working out the causal connectivities of metabolic networks from such perturbations.

10.5 The Way Ahead

The descriptions of measurement methods and data analyses make clear that much work remains to be done to obtain the full benefit of metabonomics. A key problem is that hundreds, if not thousands, of metabolites remain to be identified. Eighty percent of metabolites recognized by MS do not appear in standard libraries. As evident by the number of publications that appear on a daily basis on metabonomics and metabolomics, much effort will be spent in filling the gaps. The reader is encouraged to use this chapter as only a place marker for metabonomics and metabolomics in drug discovery and development and is encouraged to be up to date on his/her research as data acquisition platforms, software for analysis of such data, and strategies for interpretation will evolve very quickly.

David Wishart, University of Alberta, will endeavor to identify and quantify all of the small molecules in the human body at concentrations greater than 1 μM.[2] This challenge will be met with the lack of simple automated analytical strategies that can achieve this result in a reproducible and robust form. The main challenge remains the chemical complexity and heterogeneity of metabolites, the dynamic range of the measuring technique, and the throughput of the instruments and extraction protocols.[26]

When high fidelity data are gathered, standardization will be the key to acceptance by all who contribute, and render the data useable for drug discovery and development. Standards such as Systems Biology Markup Language (SBML) will allow interoperability between different models, including their metadata, and modeling packages. Ultimately, to maximize the value of this work, we will need to determine the directions, magnitude, and components of the series of multivariate influence vectors that exert a metabolic pressure on the individual. Alternatively, we will need to be able to build a probabilist system on the basis of prior knowledge and outcomes that takes metabolic observations and works forward to find the best model that is consistent with both.

This is only the beginning (Figure 10.1). Much more work lies ahead using an integrative biology approach where the three areas, genomics, proteomics, and metabonomics are combined to provide the foundation for rational drug discovery and development strategies as we move toward personalized medicine.

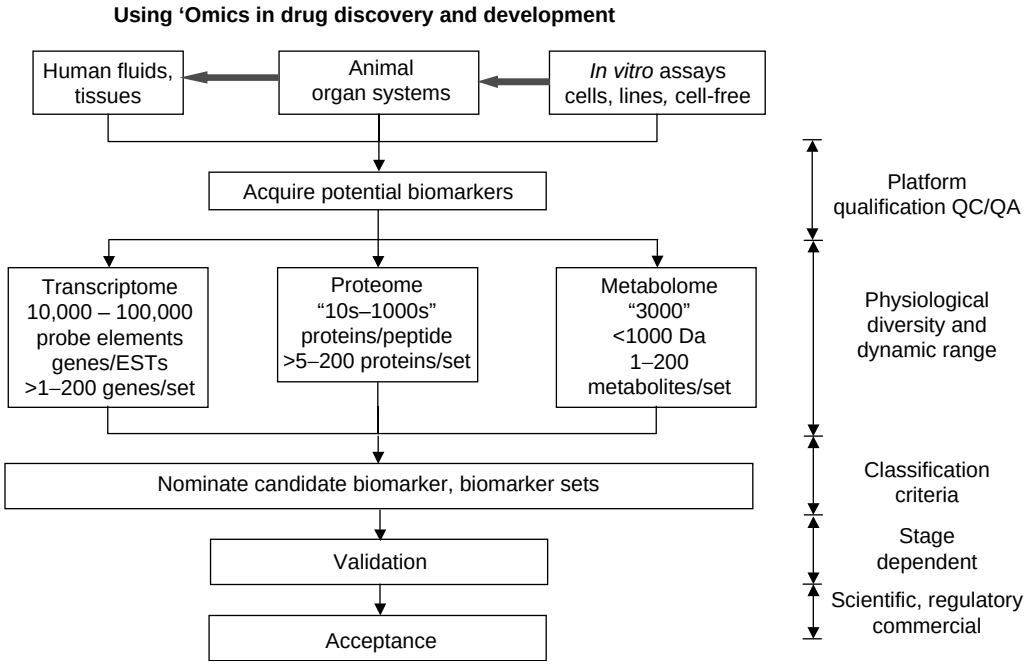


FIGURE 10.1
Characterizing the metabolome is only one component which needs to be integrated with its counterparts to realize its potential contributions to drug discovery and development. When all of the work has been done to define the universe of the transcriptome, proteome, and metabolome, and biomarker sets have been identified, 90% of the work remains in the validation and acceptance of these patterns.

References

1. http://www.metabometrix.com/metabonomics.htm#GlossaryOfTerms, Metabometrix, Ltd., 2003.
2. Hopkins, K., 'Omics: the NextGeneration, the study of metabolomics is attracting a flurry of biotechs and academics, with research implications ranging from plant biology to drug discovery, *Bio-IT World*, 48, 47–52, 2004.
3. NIH Guide METABOLOMICS TECHNOLOGY DEVELOPMENT.htm, Request for application-RM-04-002, http://www.nih.gov/, September 29, 2003.
4. Reo, N.V., NMR-based metabolomics, *Drug Chem. Toxicol.*, 25, 375, 2002.
5. Nicholson, J.K., Connelly, J., Lindon, J.C., and Holmes, E., Metabonomics: a platform for studying drug toxicity and gene function, *Nat. Rev. Drug Discovery*, 1, 153, 2002.
6. Henry, C.M., New 'Ome' in town, *Chem. Eng. News*, 80, 66, 2002.
7. Nicholson, J.K. and Wilson, I.D., Understanding 'Global' Systems Biology: metabonomics and the continuum of metabolism, *Nat. Rev.*, 2, 668, 2003.
8. Schwetz, B.A., Toxicologic pathology: looking ahead, *Toxicol. Pathol.*, 31(suppl.), 2, 2003.
9. Kettmer, K. and Hammock, B.D., Metabolomics — a new exciting field within the "omics" sciences, *Environ. Health Perspect.*, 112, 396, 2004.
10. Kell, D.B., Metabolomics and systems biology: making sense of the soup, *Curr. Opin. Microbiol.*, 7, 296, 2004.
11. Guttman, A., Multidimensional separations in the pharmaceutical arena, *Drug Discovery Today*, 9, 136, 2004.

12. Murphy, R.E., Schure, M.R., and Foley, J.P., One- and two-dimensional chromatographic analysis of alcohol ethoxylates, *Anal. Chem.*, 70, 4353, 1998.
13. Giddings, J.C., Sample dimensionality: a predictor of order-disorder in component peak distribution in multidimensional separation, *J. Chromatogr. A*, 703, 3, 1995.
14. Fiehn, O., Metabolomics — the link between genotypes and phenotypes, *Plant Mol. Biol.*, 48, 155, 2002.
15. Grivet, J.-P., Delort, A.-M., and Portais, J.-C., NMR and microbiology: from physiology to metabolomics, *Biochemistry*, 85, 823, 2003.
16. Kirn, T.F., Magnetic resonance spectroscopy may hold promise in studying metabolites, tissues, *JAMA*, 261, 1103, 1989.
17. Nicholson, J.K., Metabonomics, a generic platform for the investigation of drug toxicity, altered gene function and disease processes, *Drug Discovery World*, Summer, 23–28, 2004.
18. Lindon, J.C., Nicholson, J.K., Holmes, E., Antti, H., Bollard, M.E., Keun, H., Beckonert, O., Ebbels, T.M., Reily, M.D., Robertson, D., Stevens, G.J., Luke, P., Breau, A.P., Cantor, G.H., Bible, R.H., Niederhauser, U., Senn, H., Schlotterbeck, G., Sidelmann, U.G., Laursen, S.M., Tymiak, A., Car, B.D., Lehman-McKeeman, L., Colet, J.M., Loukaci, A., and Thomas, C., Contemporary issues in toxicology, the role of metabonomics in toxicology and its evaluation by the COMET project, *Toxicol. Appl. Pharm.*, 187, 137, 2003.
19. Brecker, L., Weber, H., Griengl, H., and Ribbons, D.W., In situ proton-NMR analyses of *Escherichia coli* HB101 fermentation in $^{1}H_2O$ and in D_2O, *Microbiology*, 145, 3389, 1999.
20. Raamsdonk, L.M., Teusink, B., Broadhurst, D., Zhang, N., Hayes, A., Walsh, M.C., Berden, J.A., Brindle, K.M., Kell, D.B., Rowland, J.J., Westerho, H.V., van Dam, K., and Oliver, S.G., A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations, *Nat. Biotechnol.*, 19, 45, 2001.
21. Griffin, J.L. and Corcoran, O., High-resolution magic-angle spinning ^{13}C NMR spectroscopy of cerebral tissue, *Magn. Reson. Mater. Phys. Biol. Med.*, 18, 1352, 2005.
22. Lacey, M.E., Subramanian, R., Olson, D.L., Webb, A.G., and Sweedler, J.V., High-resolution NMR spectroscopy of sample volumes from 1 ηL to 10 mL, *Chem. Rev.*, 99, 3133, 1999.
23. Lindon, J.C., HPLC-NMR-MS: past, present and future, *Drug Discovery Today*, 8, 1021, 2003.
24. Duran, A.L., Yang, J., Wang, L., and Sumner, L.W., Metabolomics spectral formatting alignment and conversion tools (MSFACTs), *Bioinformatics*, 19, 2283, 2003.
25. Johnson, K.J., Wright, B.W., Jarman, K.H., and Synovec, R.E., High-speed peak matching algorithm for retention time alignment of gas chromatographic data for chemometric analysis, *J. Chromatogr. A*, 996, 141, 2003.
26. Goodacre, R., Vaidyanathan, S., Dunn, W.B., Harrigan, G.G., and Kell, D.B., Metabolomics by numbers: acquiring and understanding global metabolite data, *Trends Biotechnol.*, 22, 245, 2004.
27. Mangan, S. and Alon, U., Structure and function of the feed-forward loop network motif, *Proc. Natl. Acad. Sci.*, 100, 11980, 2003.
28. Fiehn, O. and Weckwerth, W., Deciphering metabolic networks, *Eur. J. Biochem.*, 270, 579, 2003.
29. Vance, W., Arkin, A., and Ross, J., Determination of causal connectivities of species in reaction networks, *Proc. Natl. Acad. Sci.*, 99, 5816, 2002.

11

*Pharmacogenetics and Pharmacogenomics in Drug Development and Regulatory Decision-Making: Report of the First FDA–PWG–PhRMA–DruSafe Workshop**

Lawrence J. Lesko, Ronald A. Salerno, Brian B. Spear, Donald C. Anderson, Timothy Anderson, Celia Brazell, Jerry Collins, Andrew Dorner, David Essayan, Baltazar Gomez-Mancilla, Joseph Hackett, Shiew-Mei Huang, Susan Ide, Joanne Killinger, John Leighton, Elizabeth Mansfield, Robert Meyer, Stephen G. Ryan, Virginia Schmith, Peter Shaw, Frank Sistare, Mark Watson, and Alexandra Worobec

CONTENTS

*From Food and Drug Administration, Center for Drug Evaluation and Research, Rockville, Maryland, Office of Clinical Pharmacology and Biopharmaceutics (Dr. Lesko, Dr. Huang), Office of Testing and Research (Dr. Collins, Dr. Sistare), Office of New Drugs (Dr. Leighton, Dr. Meyer); WorldWide Regulatory Affairs, Wyeth Research, St. Davids, Pennsylvania (Dr. Salerno); Pharmacogenetics, Abbott Laboratories, Abbott Park, Illinois (Dr. Spear); Pharmacia Corporation, Kalamazoo, Michigan (Dr. D. Anderson, Dr. Gomez-Mancilla); Pfizer, Inc., Groton, Connecticut (Dr. T. Anderson); Genetics Research, GlaxoSmithKline, Greenford, Middlesex, England (Dr. Brazell); Wyeth Research, Andover, Massachusetts (Dr. Dorner); Wyeth Research, Chazy, New York (Dr. Killinger); Food and Drug Administration, Center for Biologics Evaluation and Research, Office of Therapeutics, Division of Clinical Trial Design and Analysis, Rockville, Maryland (Dr. Essayan, Dr. Worobec); Food and Drug Administration, Center for Devices and Radiological Health, Office of *In Vitro* Diagnostic Device Evaluation and Safety, Rockville, Maryland (Dr. Hackett, Dr. Mansfield); National Institute of Health, NHGRI, and Novartis, Gaithersburg, Maryland (Dr. Ide); AstraZeneca Pharmaceuticals, Wilmington, Delaware (Dr. Ryan); GlaxoSmithKline, Research Triangle Park, North Carolina (Dr. Schmith); Pharmacogenomics and Human Genetics, Bristol-Meyers Squibb, Princeton, New Jersey (Dr. Shaw); and Clinical Genomics, Merck and Co., Inc., West Point, Pennsylvania (Dr. Watson).

(Originally published by Lesko L.J., Salerno R.A., Spear B.B. et al., as Pharmacogenetics and pharmacogenomics in drug development and regulatory decision-making: Report of the First FDA–PWG–PhRMA–DruSafe Workshop. *J. Clin. Pharmacol.*, 43, 342–358, 2003.)

11.1 Introduction

A workshop was held in Rockville, Maryland, on May 16–17, 2002, under the sponsorship of the U.S. Food and Drug Administration (FDA), the Pharmacogenetics Working Group (PWG), the Pharmaceutical Research and Manufacturers of America (PhRMA), and the PhRMA Preclinical Safety Committee (DruSafe), the latter consisting of members of pharmaceutical companies actively engaged in pharmacogenetics.[†] This was the first agency–industry workshop on pharmacogenetics and pharmacogenomics. The goal of the workshop was to discuss the following:

1. Use of genomic technology in nonclinical and clinical drug development
2. Issues, limitations, and questions related to the application of pharmacogenetics and pharmacogenomics
3. Future direction of regulatory policy and guidances for industry

The purpose of this report is to summarize the key ideas and recommendations that were identified and discussed at the workshop. There has been no attempt to summarize each individual plenary lecture; rather, the report is organized by key topic areas that emerged during the workshop.

11.2 Definitions

There is a diversity of opinion regarding definitions and benefits of pharmacogenetics and pharmacogenomics.[1–3] For example, pharmacogenetics is often considered to be the study of inter-individual variations in DNA sequence related to drug absorption and disposition (pharmacokinetics, PK) or drug action (pharmacodynamics, PD). Polymorphic variation in the genes that encode the functions of transporters,

[†]The workshop agenda and workbook can be found at www.fda.gov/cder/calendar/meeting/phrma52002/default.htm and www.fda.gov/cder/calendar/meeting/phrma52002/workbook.pdf, respectively.

metabolizing enzymes, receptors, and other proteins can result in individual differences in the dose–plasma concentration– response relationships for many important therapeutic agents. In contrast, pharmacogenomics is thought to be the application of genomic technologies to the study of drug discovery, pharmacological function, disposition, and therapeutic response. Whole-genome or candidate gene single-nucleotide polymorphism (SNP) maps, haplotype markers, and alterations in gene expression or inactivation represent global tools for the discovery of new drugs and genetic variations that influence drug action. This workshop did not attempt to harmonize on these definitions. For the purposes of this workshop, pharmacogenetics was used to define applications of single-gene sequences or a limited set of multiple-gene sequences, but not gene expression or genome-wide scans, to study variation in DNA sequences related to drug action and disposition. Pharmacogenomics was used to define applications of genome-wide SNP scans and genome-wide gene expression analyses to study variations that influence drug action.

11.3 Overview of Pharmacogenetics and Pharmacogenomics

An important application of pharmacogenetics and pharmacogenomics to public health is the ability to determine *a priori* who will respond favorably (or unfavorably) to a given type of drug treatment. A major challenge is the inter-individual variability (population variance) and intra-individual repeatability in a clinical outcome measure (efficacy or safety) in a target population. There is substantial variability in treatment response, and data already exist to indicate that a component of variability is genetic in nature. Repeatability of a clinical response in a given subject with a chronic recurrent disease requiring continued therapy (e.g., asthma) reflects the heritability in that subject that modulates the response.[4] A challenge of future pharmacogenetic clinical trials will lie in designing the research in a way that provides information both on population variance and individual repeatability.

Pharmacogenetics and pharmacogenomics allow one to look at research with a fresh perspective (i.e., a new way to ask old questions). In this context, it is important to clearly define research goals in terms of PK, drug safety, and drug efficacy. For example, pharmacogenetics can now be used to explain (1) the well-known differences in metabolism of 6-mercaptopurine by thiopurine methyltransferase that are due to distinct population genotypes and may result in at-risk subpopulations,[5] (2) the underlying rationale (KCNE2 genotype) for trimethoprimsulfamethoxazole-induced toxicity in the form of prolongation in QTc in some individuals,[6] and (3) the interindividual variability in peak flow (efficacy) in asthmatics receiving albuterol owing to polymorphisms in the β-adrenergic receptor.[7]

Sample size requirements, as well as the role of ethnicity, will need thorough exploration for the integration of pharmacogenetics into clinical trials, depending on whether the trial goal is to identify efficacy response genotypes or exclude at-risk genotypes due to safety concerns. Study power, and thus sample size, will depend on allele frequency in the trial subjects, the effect of ethnicity on allele frequency, and the nature or type of gene action (e.g., dominant, recessive, additive, etc.).[8]

While human DNA variation as it relates to pharmacogenetic differential drug response is a static marker, RNA expression patterns as they relate to pharmacogenomic differential drug response are dynamic and change with disease state and in response to drug treatment. Therefore, expression profiling may serve as a prognostic marker of patient response on the basis of pretreatment profiles and also provide molecular biomarkers of patient response by observed changes during treatment. Differential drug responses may result from individual

heterogeneity of the molecular mechanism of disease that can be identified at the level of gene expression. RNA expression profiling can target specific genes using quantitative reverse transcriptase–polymerase chain reaction (RT–PCR) or produce a global profile using DNA chip technology. The identification of RNA patterns that correlate with patient response will allow clinicians to select patients on the basis of their predicted responses and avoid adverse reactions.

11.4 Genomic Testing and Data Quality Issues

11.4.1 Single-Nucleotide Polymorphisms Assay Validation

Various assay technologies are used to genotype SNPs or detect variant alleles in human genomic DNA samples. These include but are not limited to the following methods: (1) direct sequencing, (2) primer extension-based methods, (3) hybridization methods (including DNA "chips" of various sorts), and (4) restriction enzyme analyses. Assay validation is generally accomplished using a panel of reference samples (standards) of known genotype. The most widely accepted assay technology is bidirectional dideoxy sequencing of DNA of individual alleles. However, on occasion, sequencing may be problematic because of a high guaninecytosine (GC) content. An unresolved issue, not directly related to the technology, is evaluating the ability of a genotyping assay to identify homozygotes for rare alleles because of the difficulty in finding such individuals. The inclusion of such individuals in a reference panel may be impractical.

Haplotypes may be determined directly by sequencing individual, cloned chromosomal subsegments, or they can be derived by genotyping relevant SNPs in multigenerational families. Because these methods are considered impractical for routine application to clinical trial samples, various computational algorithms have been developed to infer haplotypes probabilistically, given the directly determined genotypes of individual, closely linked SNPs. Studies of large populations often reveal a number of less prevalent haplotypes. Comparing performance to standard assays can validate these methods. How uncommon haplotypes should be handled remains an unresolved issue.

Concerning the use of SNP and haplotype assay technology and the validation of these methods, there was consensus for a need for standardized reference materials, standards for assay validation, and specific regulatory guidance for validation criteria of the methods. There were various mechanisms suggested to address these needs, including the formation of national repositories for reference materials, specific cooperative efforts to develop standards and reference materials, and alternative strategies that could be used in the interim or in special circumstances. It was concluded that sequencing should not be the only standard that could serve as the basis for development of SNP and haplotype assays. Development of standards for use in this capacity should be a high-priority issue, although how this should be done was not delineated in the workshop. For now, *ad hoc* mechanisms such as interactions with the FDA through pre-IDE (preinvestigation of device evaluation) and pre-IND (preinvestigational new drug) meetings, use of recommendations from the American College of Medical Genetics (ACMG) and the College of American Pathologists (CAP), and the draft genetics template from the Secretary of Advisory Committee on Genetic Testing (SACGT) should be used until or unless more specific direction becomes available.

11.4.2 Reference Populations

Allele (and haplotype) frequencies typically differ among historically divergent populations. In pharmacogenetic studies, spurious associations may arise if the case (e.g., responder) and control (e.g., nonresponder) groups are drawn from genetically distinct subpopulations (e.g., based on race), even when matched for demographics or other characteristics. One approach to this problem is to test for association in cases and controls that are "genetically matched." Another is to use statistical methods based on allele frequencies at "reference" loci (e.g., in genes selected because they are unlikely to be related to drug response) to account for differences in genetic background. As pharmacogenetic data accumulate, there will be an increasing demand for population-specific data. However, there is a current need for the identification and use of reference populations, despite the inherent difficulties and ambiguities in defining such populations. In summary, the main point was that some form of defining reference population is better than none and better than each individual sponsor defining reference populations on a case-by-case basis.

11.4.3 Gene Expression Arrays

Gene expression arrays are being used nonclinically to explore associations with clinical outcomes to predict drug response, to better understand mechanisms of drug action, and to identify biomarkers of drug response that may be used to assess effectiveness or toxicity in clinical settings. The unique aspect of microarray gene expression technology is the ability to explore "genome-wide" expression in a given sample. A huge number of measurable analytes from each experiment, even with the application of 99% statistical confidence limits, could yield a significant number of false-positive and false-negative signals. Gene expression microarray technology is evolving, and its performance is not standardized among numerous platforms now available with various probes designed from different gene sequences for the same targets. The scope of information that microarrays provide requires data reduction applications to analyze, visualize, comprehend, and communicate the data output. To use these data effectively and correctly, we need reliable experimental data, sound data reduction algorithms, and publicly available biologically linked reference information. With regard to validation, each feature of an array would not necessarily need validation, but subsets of features should be evaluated with respect to performance. The development of a reference standard, such as mixed-tissue aliquots of species-specific RNA, was recognized as important for increasing confidence in this technology.

There are varying interpretations as to whether, when, or what type of microarray data on lead compounds in drug development would need to be submitted to regulatory agencies as part of an IND. The format for presentation of such data is undefined. It is not clear if the data quality is sufficiently high and convincingly reliable, given the state of the technology, for regulatory authorities to make decisions on microarray data submitted under an investigational new drug application. Many of the workshop participants suggested that all microarray data conducted under either good laboratory practice (GLP) or non-GLP conditions should be submitted to the FDA if the lead compound is covered by an IND. If microarray databases are needed for regulatory scientists to place individual sponsor microarray experimental results into proper perspective, it is not clear how this database can be built since, to date, very little data have been submitted to the agency and publicly available databases have not matured. As databases become more populated with examples, and scientific knowledge and interpretation of gene expression responses expand, there is concern that the microarray data generated today may become more

informative over time and might, therefore, affect assessments of products later during development, prior to making a final regulatory marketing approval decision. There was a strong suggestion, with general agreement, that the FDA needs to be transparent in collecting microarray data and in using it for regulatory decisions. By sharing collective knowledge gained from reviewing across applications, the FDA can improve the quality of microarray use in drug development.

Even before gene expression microarray results can be interpreted with confidence, some critical assessment of the integrity of gene expression data is required. Concerns exist about the reliability, precision, accuracy, and inter-laboratory reproducibility of data derived from global gene expression technologies. Numerous statistical, image analysis, pattern recognition, and data reduction clustering algorithms are being applied to microarray data. For screening compounds and improving understanding of drug effects on a target tissue, applying such approaches will provide a "big-picture" overview of patterns on the basis of drug class similarities but could also draw attention to the discriminating details that will distinguish among individual agents within a class. The biological interpretations, regulatory implications, and potential legal ramifications of such evaluations of product performance using global gene expression data are not well articulated at the present time.

Although there may be efforts by end users to establish standard procedures to consistently ensure and evaluate sample quality, as well as to calibrate microarrays and microarray instrumentation to ensure the integrity of their complete data sets, there is little regulatory experience to document the analytical reliability of these methods. There is a need for consensus on a standardized set of information required to fully annotate the data generated from microarray experiments. The degrees of quality control and validation that microarray manufacturers apply to their products, as well as the extent to which good manufacturing practices (GMPs) are relevant, are not always apparent. Information on manufacturing controls and postmanufacturing lot-to-lot quality control functional performance/pass–fail measures by microarray providers is important but lacking for many but not all manufacturers. There are little data from well-controlled clinical trials showing that gene expression microarray-based tests are appropriately precise, accurate, and reproducible between laboratories to be clinically valid for patient selection. Communication at the early stages of clinical development between regulatory agencies, diagnostic companies that are developing genome-based tests, and pharmaceutical companies developing drugs will enhance the possibility that test performance can be validated during clinical trials for utility in the clinical setting.

11.5 Preclinical Pharmacology, Safety, and Toxicogenomics

11.5.1 Definitions

The development of genome-based technology and subsequent genetic information has led to an entirely new field in preclinical safety and pharmacology, and toxicogenomics. As with the terms pharmacogenetics and pharmacogenomics, there are many diverse definitions of toxicogenomics.[9] For the purposes of this workshop, toxicogenomics was defined as the application of genomic concepts and technologies to the study of drug toxicity. This definition includes studies of global approaches to study alterations of gene expression that influence, predict, or help define drug toxicity.

11.5.2 Overview of Toxicogenomics in Drug Development

Toxicogenomics is applied in modern drug development in three ways as discussed below.

11.5.2.1 Predictive Toxicogenomics

This is used for screening in drug discovery and prioritization of lead compounds to improve the quality of compounds selected for development and to reduce total development time and cost.

11.5.2.2 Investigative Toxicogenomics

This is used to generate testable hypotheses. An example was presented of an application of investigative toxicogenomics to identify specific predictive biomarkers and better understand mechanisms of drug-induced vasculitis, which is a major safety concern raised by histopathology findings in animal studies with a variety of drugs. Gene expression markers of an acute-phase response, normally considered a liver response, were seen in isolated mesenteric arteries from a rat model of fenoldopam-induced vasculitis. However, further studies are needed to identify specific biomarkers of vasculitis that will define human relevance and clinical significance.

11.5.2.3 Mechanistic Toxicogenomics

This is used to improve human risk assessment by expanding accessible interspecies biomarkers of toxicity and improving the understanding of interspecies similarities and differences. A case example was presented involving identification of gene expression alterations in the rat lens associated with formation of cataracts, following administration of a 5-lipoxygenase inhibitor. The pattern of gene expression confirmed that the compound inhibited the synthesis of cholesterol and that lens crystalline structural proteins were targets of drug cataractogenesis. Further studies are needed to investigate interspecies relevance of the mechanistic-based markers of the onset and progression of cataracts.

These examples provide evidence that toxicogenomics can assist with investigations of toxicological mechanisms that involve generating and testing hypotheses, as well as identifying biomarkers useful for making interspecies predictions. However, it may not be appropriate for toxicogenomics to be used in a definitive manner at present to disprove hypotheses, confirm causality, predict safety, or replace any existing traditional safety assays. Definitive follow-up studies are needed to fully evaluate and confirm the mechanistic insights generated by toxicogenomics. Applications of the technology, formation of strategic partnerships, peer-reviewed publication of toxicogenomic data, and clarification of regulatory and industry uncertainties represent ways to advance the usefulness of toxicogenomic technology.

11.5.3 Toxicogenomics to Predict Potential Human Toxicity

A major potential use of toxicogenomics is as a more efficient alternative to traditional preclinical toxicology studies, with an added value of providing a more mechanistic understanding of different types of toxicity. The assumption is that each chemical entity acts through particular mechanisms of action that will induce, either directly or indirectly, a unique and diagnostic gene expression profile under a given set of conditions.[10] In some cases, changes in early gene expression result in pathological outcomes.[11] Pilot studies in lower eukaryotes and mammals have demonstrated that it is possible to identify common gene expression profiles of drugs with similar mechanisms of action. For example, Waring et al.[12] were able to use gene expression profiling to cluster hepatotoxins based on their

mechanism of toxicity. In addition, with the growing availability of DNA chips from several different species, it may become possible for toxicity DNA expression measures to be extrapolated across at least some species, thus providing insight into more appropriate species selection for long-term toxicology studies.

However, the utility of toxicogenomics in replacing traditional preclinical toxicology studies is controversial, with many believing that gene expression profiling should be used for hypothesis generation rather than for predicting or confirming toxicity or the lack of toxicity. Here the concerns are as follows:

1. That there are many different effects of toxicants aside from changes in gene expression, such as effects on membrane integrity, and these effects might not be extrapolated from single endpoint gene expression data.
2. In tissues with diverse cell populations, critical gene expression alterations in minor cell populations, which will deleteriously affect organ or tissue function, may be missed.
3. Without functional knowledge of what observable alterations in gene expression might mean, it is difficult to associate such changes with toxicity. To do so, there needs to be an understanding of how the molecular changes are manifest at the cellular and tissue levels, including correlation of changes of gene expression with changes in protein expression. As a result, there is a need for a multidisciplinary approach to understanding mechanistic toxicology.[13]

11.5.4 Routine Use of Toxicogenomic Data for Making Safety Decisions

There were many diverse opinions about the routine use of genomic data in toxicology studies and whether such studies should be conducted under GLP conditions. On the one hand, many participants suggested that genomic data from animal studies are not sufficiently understood to be predictive of human toxicity and should not be routinely collected in GLP toxicity studies. In contrast to traditional toxicology data, it is not yet certain how to interpret genomic data, particularly when they are collected under various nonstandardized experimental conditions, at different points in time after dosing, and by using diverse genomic technologies. The value of non-GLP studies, however, is that they may be exploratory of drug safety or even hypothesis generating and, as such, might not need to be reported to the FDA under an IND. On the other hand, others suggested that all genomic data from animal studies with lead compounds should be part of the preclinical safety database, especially if these data provide biomarkers associated with pathology data, and should be integrated into the overall safety evaluation of a new molecular entity (NME). In some cases, expression microarray data will give a snapshot of activity in many different biological pathways of importance, but in terms of risk assessment, the specific pathways would have to be well understood in high functional detail to be predictive of human toxicity. This is especially true because gene expression changes do not necessarily correlate with changes in protein expression. In a few other cases, genomic data may be considered reportable as strong evidence of human adverse events. Currently, the FDA has not seen many preclinical genomic expression profiles in new drug applications, even though these genomic data may be regarded as toxicology or possibly pharmacology data that are required to be submitted to the agency.

11.5.5 Value of Toxicogenomic Data to Industry and Regulatory Authorities

Most workshop participants agreed that toxicogenomic data have their greatest utility in preclinical mechanistic studies when used in conjunction with other global "-omics"

assessments to generate hypotheses and provide possible explanations of observable toxicity. The transition of toxicogenomic data from mechanistic to predictive will be an evolutionary process and not a revolutionary one. It is important for industry and regulatory agencies to work together to advance the science by sharing data and exchanging knowledge.

11.5.6 Use of Genomic Microarray Data with Standard, Short-Term Toxicology Studies to Guide Study Design or Species Selection for Long-Term Toxicology Studies

Workshop participants concluded that results from genomic studies are not mature enough to have predictive utility, especially when considering species selection for long-term toxicology studies. The main concerns were that genomic approaches are primarily used to evaluate toxicity in selected tissues, primarily liver and kidney. In addition, whereas the human, rat, and mouse genomes are fairly well characterized, the genomes for monkey or dog, two species that are vital in traditional preclinical safety studies, are not. However, genomic technologies may be useful in understanding the genomic makeup of "new species" such as in animal knockout models and may eventually provide a scientific rationale for not selecting rats or dogs under certain circumstances as the conventional target species in traditional toxicology studies.

11.5.7 Use of Reference Databases to Interpret Toxicogenomic Results and Predict Potential Human Toxicity

A contemporary example was presented that described the building of a reference toxicogenomic database using rat and human microarrays. A large database containing gene expression data, histopathology outcomes, and hematology and clinical chemistry data was constructed from samples of rat and human hepatocytes exposed *in vitro* and from samples from livers of rats exposed to commercially available pharmaceuticals and chemical toxicants at multiple doses and time points. By applying statistical and clustering algorithms to this database, sets of multiple discriminator genes were identified to predict human drug toxicity. While sets of "common toxicology markers" had predictive power, no single dysregulated gene demonstrated the power to predict toxicity. The database and modeling algorithms were then used to assess the gene expression changes observed at 6 and 24 h postdose in livers from rats exposed to three acetylcholinesterase inhibitors. Doses of each chemical were chosen to achieve similar numbers of total expression changes but without observable liver pathology. A model was used to predict human liver toxicity for one but not for the other two inhibitors. Subsequently, using 18 blinded sets of 24-h gene expression data that were generated by 7 different laboratories, 15 of 18 sets were accurately identified ("predicted") as being derived from animals exposed to known liver toxicants, peroxisome proliferators, enzyme inducers, or vehicle. However, the ability to confidently make such assessments using data generated at different sites can be compromised if the data are not compared with each other or with information in the reference database. Thus, an assessment of a set of "invariant genes" offered potential for standardization among experiments. Standards were under development to harmonize different datasets. Preliminary data comparing common sets of compounds *in vitro* using primary rat hepatocytes and *in vivo* using rat liver samples supported the working hypothesis that coupling *in vitro* models with toxicogenomics was advantageous. This approach may then be used to screen and help pharmaceutical sponsors to prioritize drugs being considered for clinical development by helping predict and avoid compounds that may be more likely to lead to specific types of human toxicity.

Most workshop participants concluded that while much progress using internal, proprietary, or public toxicogenomic databases has been made, such databases are not yet evolved to be fully predictive of human toxicity. A key issue is how to interpret these databases. Many also cautioned that exploratory mining and cluster analyses of this database may digress from the basic biology of the chemical or drug because there are many different ways to query a database and arrange a clustering algorithm output.

11.5.8 Guidance for Industry on Toxicogenomics

The workshop participants concluded that it is premature for the FDA to write a guidance for industry because the science of genomics and its applicability to preclinical toxicogenomics assessment is rapidly evolving. If a guidance were to be developed in the future, questions and concerns would be raised about the following:

1. The purpose of the guidance
2. Recommendations on the format and analytical strategy for submitting genomic data
3. The process to be followed and the action taken if a gene expression pattern seems to be linked to a toxic event
4. Whether a specific type of microarray platform for certain experimental conditions would be recommended
5. Whether a guidance would compel a company to engage in genomics research even if it was not prepared to do so

11.6 Pharmacogenetics and Pharmacogenomics in Early Clinical Development (Phase I/II Trials)

11.6.1 Background

Pharmacogenetics and pharmacogenomics are beginning to be integrated more into early clinical development programs and as important components in plans to achieve the overall goals of Phase I and II trials, such as safety, tolerability, pharmacokinetics/pharmacodynamics (PK/PD), dose ranging, drug–drug interactions, and potentially proof of concept for efficacy hypotheses.

Genetic variants have been identified in gene coding for proteins affecting the processes leading up to a drug response. For several drug-metabolizing enzymes, as well as some drug targets or related receptors, the frequency and functional significance of these variants have been explored in clinical studies. For most genes, the functional consequences of genetic variation remain poorly characterized or relatively unknown.

Candidate gene approaches, in which there is a sufficient hypothesis, are often used in early-phase clinical trials. Because drug response is likely to involve variants from multiple genes and from genes not previously hypothesized to be involved in drug response, an alternative to a candidate gene approach is the use of "unbiased" or "hypothesis-generating" full-genome scans using an SNP map. This approach remains experimental and may require larger subject numbers, but because of the larger number of markers examined, this approach could increase the cost, time, and efforts of the study. However, it is anticipated that technological advances will make these studies more

affordable in the future. Such proof-of-concept (PoC) research also involves significant use of novel biostatistical techniques in the assessment of linkage disequilibrium, haplotype maps, and the identification of informative SNP sets.

The number of controlled clinical studies seeking to identify and validate protein or RNA expression profiles as prognostic markers of drug response (e.g., cDNA microarray analyses of primary breast tumors and prognosis) lags behind traditional pharmacogenetics. For example, an informal survey by the FDA examined more than 70 INDs and NDAs that integrated pharmacogenetic and pharmacogenomic tests into early-phase development. Eighty percent of these applications were related to cytochrome P450 (CYP) DNA variants affecting drug metabolism. There were no examples of expression profiling identified in this survey.

11.6.2 Rationale for Use of Pharmacogenetics and Pharmacogenomics

A major goal of pharmacogenetic and pharmacogenomic analyses in early-phase clinical studies is to identify subpopulations of subjects with an improved safety and efficacy profile. However, there are various views on how pharmacogenetics and pharmacogenomics could be used in the design of clinical trials. Some examples are as follows:

- Some believe that once a genotype or mRNA expression profile demonstrates a relationship with a phenotype of potential clinical importance, inclusion/exclusion criteria based on this association should be added to future studies. Others believe it is important to validate the association by replication prior to selecting patients on the basis of genotype or phenotype.
- Some believe subjects potentially at risk for adverse events or nonresponse based on prior pharmacogenetic and pharmacogenomic studies could be excluded from future studies. Others believe that because these at-risk patients may receive the drug in the real-world setting, they should be included but possibly studied in a closely monitored setting.

Various approaches have been used to determine when blood samples should be collected for pharmacogenetic and pharmacogenomic research, ranging from collecting samples in all early-phase studies to collecting samples in only those studies with narrowly defined and limited hypotheses. Another approach is to collect samples in certain types of studies (e.g., drug interaction studies) or in studies from certain development phases only (e.g., Phases I, II, or III). Some sponsors and contract clinical research organizations routinely collect blood to screen their volunteer panels to determine their genotype for important metabolic enzymes with known polymorphisms such as CYP2D6.

11.6.3 When Is It Appropriate to Use Pharmacogenetics and Pharmacogenomics for Inclusion/Exclusion Criteria (or Stratification), or When Is It Appropriate for Pharmacogenetic and Pharmacogenomic Relationships to be Explored *Post Hoc*?

Participants concluded that the use of pharmacogenetics and pharmacogenomics in early Phase I would be an important step in generating information that could be confirmed or validated in Phase II. Most participants thought that pharmacogenetic and pharmacogenomic objectives in Phase I should be considered exploratory (not confirmatory) in nature, and that pharmacogenetics and pharmacogenomics should not be used as inclusion/exclusion criteria for a single- or multiple-dose first-time-in-humans (FTIH) study. However, under a few circumstances for variants with well-established functional significance (e.g., alleles of CYP2D6, CYP2C19), a Phase I study may be confirmatory and

still use small numbers of patients. Also, in drug interaction studies for drugs metabolized by polymorphic CYP enzymes (e.g., CYP2D6, CYP2C9, and CYP2C19), pharmacogenetics and pharmacogenomics should be considered as inclusion/exclusion criteria or for stratification.

Because of the availability of plasma concentration–time data following administration of many doses, Phase I provides a unique opportunity to explore the relationships between genetic variants in genes related to metabolic enzymes and transporters and the PK properties of the compound. Because Phase I dose-ranging studies include a broader range of doses of drugs administered to subjects (and possibly resulting in a higher rate of adverse events), and since some of these doses are rarely repeated in later trials, Phase I provides a unique opportunity to explore the relationships among genetic variants in the drug target and adverse events. In addition, when an appropriate tissue is available, Phase I provides an early opportunity to evaluate gene expression profiles to identify associations with safety and efficacy at a wide range of doses.

When pharmacogenetics and pharmacogenomics are included in Phase I, gene variants that should be studied should include those encoding for the activity of drug-metabolizing enzymes and transporters, the drug target, and any pharmacological pathways related to important safety outcomes in healthy volunteers. In addition, if patients are studied, validated disease genes (that could affect drug response) should be evaluated. Investigators should attempt to identify trends and how well the clinical data correspond to preclinical data.

With regard to Phase Ib/2a (dose-ranging and PoC studies), there is a need to have some confirmation or validation before using pharmacogenetics and pharmacogenomics as inclusion/exclusion criteria or for stratification in a Phase Ib/2a study. There could be a benefit to include all genotypes in PoC studies if all patients have a potential to benefit. However, the dose may require adjustment for subjects with genotypes resulting in reduced enzyme activity (e.g., CYP2D6).

Several participants felt that pharmacogenetics and pharmacogenomics should not be treated differently from any other covariate. In some cases, more confirmation would be required prior to using these data as inclusion/exclusion criteria (or for stratification), and pharmacogenetics and pharmacogenomics would be viewed as a covariate in the *post hoc* analyses.

In this regard, the following factors should be considered in the context of stratification or use as inclusion/exclusion criteria.

11.6.3.1 Therapeutic Area

For life-threatening indications such as oncology, many participants felt that there is more willingness in clinical practice to stratify on the basis of pharmacogenetics and pharmacogenomics. In other therapeutic areas in which many effective agents are already available and physicians are accustomed to titrating the dose in individual patients (e.g., depression), it was thought that there was less willingness to stratify on the basis of pharmacogenetics and pharmacogenomics until differences in response are linked to genotype differences. For many other therapeutic areas (e.g., asthma and respiratory), there is some interest in using pharmacogenetics and pharmacogenomics for stratification.

11.6.3.2 Safety or Efficacy

Many participants expressed that there is more willingness to stratify on the basis of pharmacogenetics and pharmacogenomics for safety than for efficacy. For example, it was felt that for a CYP2D6 substrate (confirmed by *in vitro* and Phase I data), a stratified

design in Phase II, with CYP2D6 poor metabolizers (PMs) being randomized to a standard or low dose, was appropriate. The seriousness and consequences of nonresponse or an adverse event and the ethical implications of these outcomes need to be considered. Using pharmacogenetics and pharmacogenomics was suggested as being analogous to the study of patients with reduced renal function. In the case of low creatinine clearance, patients may be excluded from Phase II/3 studies until a small study in the renally impaired population is conducted. Then the Phase II/3 studies may be amended to include such patients, or these patients may continue to be excluded, depending on results of the small study.

11.6.3.2.1 Magnitude of Effect Relative to the Therapeutic Index

If a drug has a narrow therapeutic index, it may be appropriate to stratify on the basis of pharmacogenetics and pharmacogenomics. If the therapeutic index is wide, some large effects may not be important enough to warrant stratification.

11.6.3.3 Stage of Knowledge of the Variants or Expression Profile

For gene variants with known functional significance, less information may be required prior to using pharmacogenetics and pharmacogenomics as inclusion/exclusion criteria or for stratification. For example, if Phase I results along with relevant *in vitro* data showed a reasonable relationship between CYP2D6 status and PK inter-individual variability, then this would be considered a valid reason to stratify or use as inclusion/exclusion criteria in Phase II. For a drug with a narrow therapeutic index, stratification should be done by CYP2D6 status and not exclude PMs when all patients can benefit, although CYP2D6 PMs may require lower doses. For variants in the drug target in which the functional consequences are unknown, one should consider "all comers" and analyze data using the genetic variant as a covariate in a *post hoc* analysis. Even after one or two Phase I studies show that a variant in the drug target may affect drug response, there may not be enough information to exclude a population from Phase II studies, unless more is understood about the functional consequences of the genetic variants.

11.6.3.4 Allele Frequency of the Variant

If the allele frequency is common (e.g., >15%), the optimal approach would be to conduct a single trial and stratify by genotype. However, if the frequency is low (e.g., <10%), it may be less feasible to evaluate both genotypes in the same study because recruitment of subjects with the minor allele would be much slower than for those with the dominant allele. Two separate trials would allow the drug to be progressed more quickly in the dominant population. However, if separate studies were conducted, there would be less information about the analytical and clinical sensitivity and specificity of a pharmacogenetic and pharmacogenomic test. There may be some cases when safety data, but not efficacy data, from a pharmacogenetic and pharmacogenomic subgroup may be studied or other cases when a dose-ranging study in the minor pharmacogenetic and pharmacogenomic subgroup may be conducted postmarketing. One should not assume that PMs should be treated the same for all substrates (i.e., there is a need for case-by-case review).

11.6.3.5 Dose Response

If there were a rationale for pharmacogenetics and pharmacogenomics to affect dose response, then the optimal approach would be to evaluate dose response in both groups (e.g., CYP2D6 extensive and poor metabolizers).

11.6.3.6 Other Factors

These include factors that need to be considered such as (1) biological validity of results (i.e., are the results consistent with theoretical or *in vitro* data?), (2) the extent to which they have been replicated, (3) the number of candidate genes or SNPs affecting the phenotype, (4) the need for the optimal timing and special tissue handling for RNA expression profiling, and (5) the validity of supervised machine learning programs, if used for RNA expression profiling.

11.6.4 When or How Should Samples Be Collected for Genotyping/mRNA Expression Profile/SNP Profiling?

While pharmacogenetic and pharmacogenomic information could be useful in Phase I/2 studies, the science is not at the point where samples should be collected in all studies. If there is a strong scientific rationale for obtaining pharmacogenetic and pharmacogenomic data, then the samples should be collected. However, it should be kept in mind that if the results of Phase I/2 studies show the value of sample collection, it might be too late to initiate sample banking.

Some examples of scientific rationale to perform pharmacogenetic and pharmacogenomic studies include (1) a compound metabolized by a polymorphic enzyme, (2) drug interaction studies involving any substrates with polymorphic enzymes (e.g., probe cocktail), and (3) variants in the drug target or pathways known to affect safety (e.g., long QT genes).

If the metabolism of a compound is not fully understood during FTIH studies, pharmacogenetic and pharmacogenomic samples should be collected from early Phase I studies to provide the ability to examine genotype-outcome associations that may be pertinent in the later development phases.

Although DNA sample collection is independent of time, RNA samples require critical timing of the sample and more time-critical tissue processing. Thus, DNA collection may be easier to justify scientifically than RNA. Some believe RNA sample collection should only be done when there are preclinical data to suggest optimal timing. Others believe that optimal timing in animals in preclinical studies might not be predictive of optimal timing in humans; therefore, the stringency for RNA should be no different than for DNA. With RNA, one must consider what tissue is available, which for most studies is linked to blood, skin, or a pathological tissue biopsy.

11.6.5 How Will Preknowledge of Genetic Susceptibility to Pharmacologically Predictable Adverse Events or Nonresponse Obtained in Early-Phase Development Affect the Risk/Benefit Assessment and Product Labeling?

If a pharmacogenetic and pharmacogenomic subgroup had improved efficacy or safety, how the information is included in the label would be dependent on the risk/benefit assessment, with life-threatening events being considered different from less severe adverse events. The pharmacogenetic and pharmacogenomic relationship may be described under "Clinical Pharmacology," "Indications and Usage," "Warnings or Precautions," or "Dosage and Administration." The label would inform clinicians that there is a genotype or phenotype test available, but it may not have to be done prior to dosing unless warranted. If the diagnostic test must be done prior to dosing, then it may be stated under "Indications and Usage" (e.g., approved labeling for Herceptin® [trastuzumab]).

The decision of whether a pharmacogenetic and pharmacogenomic test is necessary prior to dosing will be dependent on many factors, including the following: (1) if safety, the seriousness of the adverse event; (2) if efficacy, the consequences of nonresponse; (3) the incidence of the clinical outcome; (4) the variability in the clearance of the drug; (5) how well an adverse event can be managed (i.e., if it can be recognized easily without a genetic test and whether toxicity is reversible); (6) need for education of physicians and third-party payers; and (7) feasibility of accessing and using the test in clinical practice. For example, oncologists would be more likely to use a pharmacogenetic and pharmacogenomic test prior to treatment if it would improve efficacy or safety. On the other hand, an allergist who has a patient with allergies may want immediate relief for the patient and is unlikely to wait several days to a week before writing a prescription so that a genetic test can be run to predict whether the patient is at risk for a drug-related headache.

11.7 Pharmacogenetics and Pharmacogenomics in Late Clinical Development (Phase III Trials)

11.7.1 Background

Clinical studies intended to evaluate the safety and efficacy of new drugs in development generally involve large number of patients and are critical to the evaluation and approval of a new drug. The role of pharmacogenetics in late clinical development (Phase III trials) is to focus on either the further exploration of genetically defined populations or the confirmation of pharmacogenetic data from these populations to support efficacy, safety, and the labeling of the drug. When subsets of patients respond to a drug differently, clinical trials can be designed to take advantage of these differences. For instance, trials could be limited to individuals who are more likely to receive a clinical benefit or less likely to suffer an adverse response. However, in so doing, the trials may not adequately determine the safety and efficacy of the drug in all individuals who might be exposed to it. Approaches to having the study power for genotypes need to be reconciled to take into account the need for a thorough assessment of the beneficial and harmful effects of the drug once it is in clinical practice.

In another potential use of pharmacogenetics, drugs that have not been shown to be adequately safe and effective in a clinical trial on an entire population may achieve that goal in a genetically defined subset of the population. Because genotypes do not change in an individual, it should be possible to detect a group who will derive a clear clinical benefit by reanalyzing the data from a previously completed trial through genetic stratification. In this way, a drug that is otherwise unregisterable might be approved for the genetically defined group. However, this use of pharmacogenetics poses a number of questions for which there are no definitive answers at present:

- Are there conditions under which such a retrospective study would be acceptable for drug registration?
- To what extent do these studies need to be replicated?
- Can the data from such studies, not specifically designed as a pharmacogenetic study or even from investigative studies with no genetic hypothesis, be used for registration?
- What constitutes acceptable data in such studies?
- How do these data apply from one racial or ethnic group to another when there may be significant differences in allele frequencies between groups?

To date, there appears to have been relatively little application of pharmacogenetics in Phase III studies and subsequently in regulatory decision-making. Few examples exist that can be used to assess various models for pharmacogenetic trials. The workshop focused discussion on the types of trials that might be conducted and to estimate the likely reception that such trials might receive. In practice, any such trial, especially those with novel pharmacogenetic approaches, should be discussed in detail between the sponsors and the regulatory authorities and would likely be evaluated on a case-by-case basis. Only after numerous examples exist will it be possible to develop a general guidance for industry that will provide recommendations to direct the conduct of these trials.

11.7.2 How Will Conducting a Clinical Trial in a Pharmacogenetically Defined Subset of Patients Influence the Collection of Adequate Safety and Efficacy Data Prior to Registration?

Although the integration of pharmacogenetics into clinical trials is based on newer technologies and newly discovered knowledge of the genome, the issues raised by using pharmacogenetic information in selecting patients is very similar to the issues raised by other forms of enrichment. The utility of pharmacogenetic data will depend on the following:

- The robustness of the study results (i.e., how well established is the association between the pharmacogenetic enrichment biomarkers, drug exposure, and clinical endpoints?)
- Whether patients can be readily identifiable in a practice setting (i.e., can they be preidentified with readily available tests or assessments?)
- Whether there is an expectation that the drug will be used only in this enriched population in practice?

In many cases, pharmacogenetically defined patient groups will not display a dichotomous relationship between their genetic status and their response to treatment but rather will show a gradation of responses. The smaller the difference in response (efficacy or safety) between the genetically defined group and the general population, the more important it becomes to compare the response in patients who are positive for the genetic biomarkers and negative for the genetic biomarkers. This will be necessary not only to help establish the clinical utility of the biomarkers but also to establish an overall risk/benefit ratio for the treatment if used in the general population, including those negative for the genetic biomarkers. If there is reasonable expectation that drug use would occur in the wider population (with or without knowledge of the genetic status of the patient), or if availability of the relevant pharmacogenetic test may be limited, preapproval testing in the negative population will be necessary to ensure that the overall risk/benefit of the drug in the general population is acceptable.

The amount of clinical data needed to confirm the clinical value of a pharmacogenetic biomarker will differ depending on the prior knowledge of the genetics involved and the mechanistic understanding of the way the drug therapy works in relationship to the genetics. For example, polymorphisms in a receptor that is understood to affect drug binding will require less data for confirmation than will polymorphisms in genes whose biological role is unknown.

In most cases, there appears to be no overarching ethical reason to exclude certain subsets of patients from pharmacogenetic-based clinical trials, even those who may be at increased risks of a particular toxicity. However, each situation would need to be considered

in context, and the decision to include such patients would depend on the severity of the disease being treated, the severity and the reversibility of the known or anticipated toxicities, and the current strength of evidence on the predictive value of the pharmacogenetic assessment for that treatment and disease.

Finally, for regulatory authorities to approve a drug only for a defined pharmacogenetic subset of patients, especially if tested only in that subset for safety and efficacy, it is highly likely that a clinical diagnostic assay should be available at the time of approval of the drug. Although ideally this would be an approved *in vitro* diagnostic kit, this is not an absolute necessity because many hospital or laboratory tests are not kits but are developed and validated within the testing laboratory ("home brews"). In situations when regulatory agencies will not formally regulate the test, it is advisable to involve experts such as the College of American Pathologists while considering test standardization and other quality control aspects.

11.7.3 Under What Circumstances Can a "Pharmacogenetic Clinical Trial" Be Conducted Using Samples or Clinical Data from a Previously Completed Drug Clinical Study?

Discussion at the workshop focused on whether the results from a pharmacogenetic study can be used for drug evaluation and registration if the clinical data came from a previously completed study not originally designed for genetic stratification. Such a study would be both retrospective with respect to the collection of samples and clinical data and prospective with respect to testing a genetic hypothesis. Initially, there were two contrasting views. On the one hand, some maintained that any trial that had already been unblinded was a retrospective trial and so would not be acceptable as confirmatory evidence for regulatory approval. However, it would be an acceptable hypothesis-generating trial. On the other hand, others viewed that such a trial could have its own design, separate from the original study protocol, and that the analysis would be thoroughly blinded with respect to the clinical outcomes and so should be an acceptable study for drug registration.

Specific considerations that emerged from this discussion are the following:

- While the original study would be adequately powered for the expected outcome, the genetic study might be underpowered. However, it was pointed out that, depending on the degree of association of the genotype with the response, power in the genetic study could be sufficient. Statistical power in the genetic study could be determined at the time of protocol design if there was sufficient preliminary knowledge of the association.
- Caution is needed because the original study may have been designed for a particular population, but the genetic study comprises a restricted, different population such that the randomization may not have been considered appropriate for the genetic study.

Another key factor with regard to the acceptance of a genotype study is the biology of the gene used as a marker for stratification. In cases in which the genetic biomarker is plausibly linked with the response of interest (e.g., a polymorphism in the drug target or a drug-metabolizing enzyme), the retrospective–prospective trial may provide data that would be useful in the evaluation of the drug. Nevertheless, it was generally accepted that, under any circumstances, an independent prospective trial with genotyping included in the basic study design would be necessary. The requirement for single or multiple confirmatory trials could differ depending on whether the relevant outcome was efficacy or safety and the strength of the association.

Other important issues that need to be considered are the following:

- Careful collection and storage of the DNA samples would be necessary, both from a stability standpoint (generally not a problem) and for tracking and inventory.
- It is also crucial that the samples be collected with adequate informed consent for whatever genotyping may later take place.
- The nature of the test to be used to genotype patients in the trial and subsequently in general use is also critical.

This is especially so in tests that might involve multiple sequences or multiple genes.

There was general agreement in the workshop that a trial involving genotypic stratification using clinical data from a completed study might be considered as evidence for regulatory purposes if it were to meet several specific conditions or criteria. Those conditions would involve having (1) adequate power in the genetically defined subsets, (2) appropriate informed consent, (3) proper sample handling procedures, (4) an adequately validated genotyping test, (5) clear biological relation between the response and the gene(s), (6) a prospective hypothesis for assessing the response–genotype relationship, and (7) follow-up by an independent prospective trial. In the absence of these conditions, the prospective–retrospective trial would be useful only for purposes of hypothesis generation. Given the complexities of the possible study designs and genetic associations, the determination that a retrospectively genotyped study would be acceptable for regulatory purposes should be considered on a case-by-case basis.

11.7.4 Is It Appropriate or Possible to Use Anonymized Samples or Data in Registration Studies?

Several processes are currently being used during drug development for the collection of DNA samples and associated patient data. One recent article[14] and one regulatory guidance[15] summarize and describe terminology and processes for the collection of samples and data. Concerns exist that genetic data might be used discriminatorily, for instance, for employment or insurance. Therefore, procedures for collection and data generation have been developed to provide additional confidentiality and privacy to patients by dissociating genetic data from patient identifiers. The workshop discussion focused on the merits of two types of pharmacogenetics data: (1) data that can be linked back to a patient's code number and thus to a patient identity (coded, single-coded, deidentified, double-coded) and (2) data that cannot be linked back to a patient's record (anonymized, anonymous).

From the discussions, several important points were highlighted:

- From a regulatory standpoint, it was emphasized that data to be used for registrational purposes require an audit trail back to the medical record, as is standard for all other data submitted for this purpose. Anonymous or anonymized data[14] are not auditable and would not be appropriate for registrational use.
- Eliminating the link between samples or data and the patient record provides additional privacy and confidentiality as long as there is strict adherence to standard operating procedures that prevent matching of multiple recorded clinical phenotypes to reidentify an individual. Such samples are not auditable but are useful for research purposes and might be especially suited for hypothesis generation.

- Anonymized or anonymous data and samples can be collected in registrational studies but cannot be used to support the primary objective of the trial in which they were collected.
- Samples without a link to the patient record cannot be used to validate and support new results and hypotheses discovered during later stage development of a drug, for example, in validating the relationship between a diagnostic reagent and clinical response observed in subsequent trials or during postregistrational use.

Most workshop participants recommended that data to be used to support drug safety and efficacy for registration should have a link to the patient record. In addition, samples linked to patients' data have potential for use in subsequent validation studies that are unanticipated at the outset of drug development.

11.7.5 What Characteristics of Association Data Are Expected in Exploratory or Registration Studies?

Until there is greater experience with pharmacogenetics, the participants concluded that specific requirements for the use of pharmacogenetic data in exploratory drug development or in registration trials should be defined on a case-by-case basis. Currently, there are few examples of pharmacogenetic-based drugs on which regulatory guidance criteria can be based. In general, the requirements for using predictive genetic biomarkers are similar to those for nongenetic biomarkers and should be based on equally compelling scientific concepts and arguments.

Several limitations are inherent in applying pharmacogenetic associations derived from case control datasets, even if statistically significant results are observed. Replication of pharmacogenetic associations identified in an initial study population may not always be possible in separate studies and populations. Association studies also pose several inherent statistical challenges,[16] and sponsors are unlikely to repeat large Phase III trials solely for the purpose of confirming the genetic results. Nevertheless, the confirmation of identified genetic associations is mandatory if pharmacogenetic biomarkers are to be included in drug labeling. To accomplish this goal, researchers should identify such biomarkers in smaller exploratory studies and then confirm them in larger trials in which statistical validation can be achieved. Informed consent for genetic studies and routine collection/archiving of genetic samples also should be considered, whether there is a rationale or not, early in the drug development process[17] and should be adequate for future marker identification, drug regulation, and assay development. This will allow for genetic analysis of responses that may appear only at later stages of drug development or marketing.

Although such markers may be widely employed in the design of exploratory trials to enable validation of pharmacogenetics signals and to provide "proof of concept," the specific characteristics of genetic association data for registration and labeling of a drug will depend on the context of the experimental questions and the pharmacogenetics objectives to be addressed. For example, a predictive genetic biomarker of drug response should be sufficiently common in applicable patient populations, and the degree of enrichment of response afforded should be clinically meaningful. These requirements, in relative terms, will depend on clinical factors, including the nature and severity of disease and the extent of the medical need unmet, if any. Use of genetic biomarkers to exclude individuals or populations at risk for adverse events will require rigorously validated genetic associations, especially for serious toxicity. Acceptability for false-negative results of genetic biomarkers to predict drug toxicity will be very low. Acceptability for

false positives may be higher, especially if alternative therapies are available. It was pointed out that although such biomarkers may be widely employed in the design of exploratory trials to enable validation of pharmacogenetics signals and to provide proof of concept, the regulatory requirements for application of such biomarkers in registration trials or in drug labeling are uncertain and yet to be defined. The confidence with which any pharmacogenetic biomarker can be applied in clinical studies or practice will be enhanced by supportive biologic experimental data, providing a credible scientific rationale for a pharmacogenetic hypothesis.

11.7.6 If Exploratory Pharmacogenetic Research Is Performed during the Clinical Development of a Compound Outside the Basic Clinical Study Design, under What Circumstances Would the Results of a Pharmacogenetic Analysis Warrant Reporting to a Regulatory Agency?

The FDA regulations (e.g., CFR 312.23) cover the reporting obligations for exploratory studies of this type. Pharmacogenetic data are to be considered a part of the drug development process and should not be segregated or differentially reported and, if relevant to the safe and effective use of a product, should be reported to the appropriate regulatory agencies for review. The course of action by the regulatory agencies in response to such data will depend in part on the quality of these data (i.e., the validation of the study methodology) and statistical power of any associations. Future actions may be modified in the near term by the concept of a "safe harbor" for exploratory genomic data (discussed below), when and if this concept becomes better defined and accepted as suggested by the FDA.

11.7.7 What Would Be the Implications for Ethnic Diversity in Clinical Trials?

Throughout the workshop discussion, as well as in scientific and nonscientific literature, the terms ethnicity and race are often used interchangeably. There are sensitivities surrounding these terms, and trying to get consensus on the definitions was beyond the scope of the workshop. It was recognized that both risk for disease and desirable and undesirable drug responses are variable across the human species, and the variability is dependent on both genetic and environmental factors, many of which may differ between populations.

However, the link between a particular genotype and a clinical phenotype can only be established by analysis of individuals. This information may be best uncovered through understanding genetic diversity within a given population. Observation of phenotype – genotype relationships in different "population groups" should be driven by analysis of genetic variation and clinical parameters in parallel with classical assessments that are socially and culturally acceptable. There are several recent reviews and discussions covering genetic diversity related to risk of diseases and the probability of drug response in different populations.[18,19]

11.8 Regulatory Perspectives on Pharmacogenetics and Pharmacogenomics

11.8.1 Safe Harbor

Pharmacogenetics and pharmacogenomics are being integrated into drug development by most, if not all, pharmaceutical research companies. Submission of pharmacogenetic and

pharmacogenomic data to regulatory agencies has been limited, but regulatory authorities are very interested in enabling this technology because of its potential to improve the drug development process and public health.

Regulatory agencies are encouraging pharmaceutical companies to explore and apply toxicogenomic, pharmacogenetic, and pharmacogenomic technologies in drug development and to submit such data for regulatory review. However, sponsors have concerns that genomic-based data would be acted on prematurely by regulatory authorities to interfere with or add to the cost of drug development. The FDA expressed a willingness to explore the feasibility of a "safe harbor" for genome-based data on both lead and nonlead compounds. This term was used to describe a process in which exploratory genomic-based data generated and submitted under an active IND would be submitted to the FDA but would not undergo formal regulatory review until more is known about the validity of the technology used and the appropriate interpretation of the data. An example of exploratory genomic-based data that might have tenuous or uncertain interpretation is the activation or overexpression of an oncogene in a DNA microarray assay in rat cells. The linkage of this event to human adverse events is unknown or extremely uncertain and may be valuable only for generating new hypotheses but would not be appropriate as the basis of a regulatory decision. The value of a safe-harbor toxicogenomic database would be (1) to gain a better understanding of the relationship between RNA expression biomarkers and pathology, (2) to discover integrated knowledge from data mining across submissions, and (3) to learn how to interpret different datasets over multiple technology platforms.

There are many details of a safe harbor that would have to be worked out, including the format for the safe-harbor presentation of data, the process for submitting such data, and the procedure for regulatory review.

The FDA acknowledges the importance of having a transparent process for decision-making related to the transition or bridging of genomic data from a safe harbor to a database subject to a formal regulatory review. One suggestion by the FDA was to cosponsor additional public meetings or workshops as part of a process to develop guidance on safe harbor or on pharmacogenetics and pharmacogenomics in general.

11.8.2 FDA Perspective on Genotyping and Clinical Efficacy/Safety Trials

Although genomic-based technology may be at an early stage, the issues and questions surrounding pharmacogenetics and pharmacogenomics are not necessarily new. For example, genotyping in clinical trials represents a form of mechanistic or empirical "enrichment" (i.e., a process for selecting or excluding individual patients or groups of patients for clinical trials). Regulatory agencies are quite familiar with criteria that have been used in the past for routinely enriching clinical efficacy or safety trials for drugs such as inotropic agents, topical nitrates, antiviral drugs, and antibacterials and in diseases such as hypertension, stroke, and sepsis. A well-known example of enrichment is the enrollment of women with breast cancer who overexpress the HER-2 protein in clinical trials of trastuzumab (Herceptin).

In the broadest sense, genotyping can also be used in proof-of-principle trials and for individualization and modification of dose based on genotype. Associations between genotypes and clinical outcomes can also be explored retrospectively, as was the case for abacavir,[20,21] but these are mainly exploratory and would need confirmation in a clinical trial prospectively. An important distinction was made between two types of genome-based enrichment: the first type (preferred) is when there is a well-understood, genome-based pathophysiological ability to select responders and nonresponders, and the second type is when genomic-based predictions of differences in response are

observational, but the basis for pathophysiology is not well understood. Some key points were: (1) that if a treatment cannot be limited to a genomically defined patient population, then effects in the overall population, especially safety in the nonselected patients, need to be determined to assess the true risk/benefit of the drug, and (2) confirmation of outcomes from genomically guided clinical trials almost always need repeating in a prospective trial to be persuasive. Several study designs for clinical pharmacogenomic studies were discussed, and while all were acceptable, it is important to clearly consider the objectives of the investigation (e.g., bioanalytical performance of a genome-based diagnostic test, clinical utility of a diagnostic test, safety in nonselected patients) in deciding on the design of the clinical trial.

11.8.3 European Agency for the Evaluation of Medicinal Products Perspective on Pharmacogenetics and Pharmacogenomics

It is important to consider the impact of the known pharmacogenetics (e.g., polymorphism in genes that code for drug-metabolizing enzymes) in dealing with the new pharmacogenomics in delineating a way forward. Concern was expressed about the high percentage of patients who do not respond or respond incompletely to drugs, as well as the substantial morbidity and mortality due to adverse drug reactions, and that perhaps greater attention should be paid to the known pharmacogenetics. Principles related to genes and drug response are already included in several European and international regulatory guidances. It was pointed out that the new pharmacogenomics is in a transitional phase where genome-based science is moving from research to the clinic and from exploratory to confirmatory research. Regulatory agencies are preparing themselves for the anticipated increase in submissions that include pharmacogenetics or pharmacogenomics. Within Europe, a Committee for Proprietary Medicinal Products (CPMP) expert group was formed in April 2001, and a CPMP position paper on terminology[15] was released for consultation and comment in December 2001; this position paper was adopted by the CPMP in November 2002. Two concerns were expressed in regard to *post hoc* genomic-based association studies, similar to the FDA concerns: first, the reliability and reproducibility of findings and, second, the need to confirm findings in a prospective clinical study that defines the sensitivity and specificity of the genetic marker. It was emphasized that regulatory agencies and industry must continue to maintain dialogue as the field moves forward quickly and that regulatory agencies should consider future guidances depending on the level of scientific knowledge, experience, and applicability of pharmacogenetics and pharmacogenomics.

11.9 Summary and Conclusions

The workshop concluded with a panel discussion, along with questions and answers from the audience related to future deployment of pharmacogenetics and pharmacogenomics in drug development and regulatory decision-making. The broad conclusions from the workshop are as follows:

1. Pharmacogenetics and pharmacogenomics should be considered in all phases of drug development because these sciences will only improve our understanding of the safety and efficacy of new drugs and improve the development of optimal dosing regimens. The use of genetic and genomic technologies, however, should be driven by science and applied where it can improve decision-making from lead compound selection to allowing market access.

2. Greater clarity on the most appropriate applications of pharmacogenetic and pharmacogenomic biomarkers in drug development is needed to advance the field.
3. Continued dialogue between academic researchers, industry scientists, and regulatory agencies is needed to help guide strategies for exploiting pharmacogenetic and pharmacogenomic information to optimize risk/benefit ratios.

The challenge of advancing pharmacogenetics and pharmacogenomics has many dimensions: scientific, economic, social, and political. In light of the huge potential for these sciences to improve the drug development process and address future public health needs, professionals from academia, industry, and regulatory agencies need to work together to achieve the potential that pharmacogenetics and pharmacogenomics offers to society. Given this challenge and the impact that these sciences can have as they evolve rapidly in the upcoming years, there are plans to conduct follow-up public workshops that will focus on subsets of issues identified in this workshop to develop a blueprint for a way forward.

Acknowledgments

We express our gratitude to the plenary speakers: Wayne Anderson, PhD, GlaxoSmithKline, Research Triangle Park, North Carolina; Andrew Dorner, PhD, Wyeth Research, Andover, Massachusetts; Jeffrey M. Drazen, MD, New England Journal of Medicine, Cambridge, Massachusetts; Steven Galson, MD, Center for Drug Evaluation and Research, Food and Drug Administration, Rockville, Maryland; Baltazar Gomez-Mancilla, MD, PhD, Pharmacia Corporation, Kalamazoo, Michigan; James T. Mayne, PhD, Pfizer, Inc., Groton, Connecticut; Donna Mendrick, PhD, Gene Logic, Inc., Gaithersburg, Maryland; Marisa Papluca-Amati, MD, The European Agency for the Evaluation of Medicinal Products, London, England; Brian Spear, PhD, Abbott Laboratories, Abbott Park, Illinois; and Robert Temple, MD, Center for Drug Evaluation and Research, Food and Drug Administration, Rockville, Maryland.

References

1. Lesko, L.J. and Woodcock, J., Pharmacogenomic-guided drug development, *Pharmacogenomics J.*, 2, 22–24, 2002.
2. Roses, A.D., Pharmacogenetics and the practice of medicine, *Nature*, 405, 857–865, 2000.
3. Evans, W.E. and Relling, M.V., Pharmacogenomics: translating functional genomics into rational therapeutics, *Science*, 286, 487–491, 1999.
4. Malmstrom, K., Rodriguez-Gomez, G., Guerra, J., Villaran, C., Pineiro, A., Wei, L.X., Seidenberg, B.C., and Reiss, T.F., Oral montelukast, inhaled beclomethasone, and placebo for chronic asthma: a randomized, controlled trial, *Ann. Intern. Med.*, 130, 487–495, 1999.
5. Weinshilboum, R., Thiopurine pharmacogenetics: clinical and molecular studies of thiopurine methyltransferase, *Drug Metab. Dispos.*, 29, 601–605, 2001.
6. Christiansen, M. et al., The genetics of long QT syndrome: mutation detection in 100 families, *Am. J. Hum. Genet.*, 69(Suppl.), 604, 2001.
7. Liggett, S.B., The pharmacogenetics of beta 2-adrenergic receptors-relevance to asthma, *J. Asthma*, 105, S487–S492, 1999.

8. Elston, R.C., Idury, R.M., Cardon, L.R., and Lichter, J.B., The study of candidate genes in drug trials: sample size considerations, *Stat. Med.*, 18, 741–751, 1999.
9. Aardema, M.J. and MacGregor, J.T., Toxicology and genetic toxicology in the new era of "toxicogenomics": impact of "-omics" technologies, *Mutat. Res.*, 499, 13–25, 2002.
10. Fielden, M.R. and Zacharewski, T.R., Challenges and limitations of gene expression profiling in mechanistic and predictive toxicology, *Toxicol. Sci.*, 60, 6–10, 2000.
11. Corton, J.C., et al., Entering into the era of toxicogenomics with DNA microarrays, *CIIT Activities*, 19, 1–9, 1999.
12. Waring, J.F., Jolly, R.A., Ciurlionis, R., Lum, P.Y., Praestgaard, J.T., Morfitt, D.C., Buratto, B., Roberts, C., Schadt, E., and Ulrich, R.G., Clustering of hepatotoxins based on mechanism of toxicity using gene expression profiles, *Toxicol. Appl. Pharmacol.*, 175, 28–42, 2001.
13. Pennie, W.D. et al., The principles and practice of toxicogenomics: applications and opportunities, *Toxicol. Sci.*, 54, 277–283, 2002.
14. Spear, B.B. et al., Terminology for sample collection in clinical trials, *Pharmacogenomics J.*, 1, 101–103, 2001.
15. Committee for Proprietary Medicinal Products, Position paper on terminology in pharmacogenetics, The European Agency for the Evaluation of Medicinal Products, London, 2002. (Available at http://www.emea.eu.int)
16. Ioannidis, J.P., Ntzani, E.E., Trikalinos, T.A., and Contopoulos-Ioannidis, D.G., Replication validity of genetic association studies. *Nat. Genet.*, 29, 306–309, 2001.
17. Anderson, D.C., Gomez-Mancilla, B., Spear, B.B., Barnes, D.M., Cheeseman, K., Shaw, P.M., Friedman, J., McCarthy, A., Brazell, C., Ray, S.C., McHale, D., Hashimoto, L., Sandbrink, R., Watson, M.L., Salerno, R.A., Cohen, N., and Lister, C.E., Elements of informed consent pharmacogenetic research: perspective of the pharmacogenetics working group, *Pharmacogenomics J.*, 2, 284–292, 2002.
18. Risch, N., Burchard, E., Ziv, E., and Tang, H., Categorization of humans biomedical research: genes, race and disease, *Genome Biol.*, 3, comment 2007, 2002.
19. Aldhous, P., Geneticist fears "race neutral" studies will fail groups, *Nature*, 418, 356–357, 2002.
20. Mallal, S., Nolan, D., Witt, C., Masel, G., Martin, A.M., Moore, C., Sayer, D., Castley, A., Mamotte, C., Maxwell, D., James, I., and Christiansen, F.T., Association between presence of HLA-B*5701, HLA-DR7, and HLA-DQ3 and hypersensitivity to HIV-1 reverse-transcriptase inhibitor abacavir, *Lancet*, 359, 727–732, 2002.
21. Hetherington, S., Hughes, A.R., Mosteller, M., Shortino, D., Baker, K.L., Spreen, W., Lai, E., Davies, K., Handley, A., Dow, D.J., Fling, M.E., Stocum, M., Bowman, C., Thurmond, L.M., and Roses, A.D., Genetic variations in HLA-B region and hypersensitivity reactions to abacavir, *Lancet*, 359, 1121–1122, 2002.

12

Drugs from Molecular Targets for CNS and Neurodegenerative Diseases

William T. Comer and Gönül Veliçelebi

CONTENTS

12.1 Drug Discovery

The discovery of a drug starts with a concept of the pathogenesis of a disease, and in some cases, includes a hypothesis of the molecular mechanism underlying the pathophysiology. The principal components of a drug discovery program include (1) a sensitive, rapid, primary screening assay for evaluating large numbers of compounds for a specific biologic end point; (2) secondary assays to assess in more detail the biological effectiveness of the compounds and any cross-reactivity with other pathways; and (3) a reliable animal model that reflects the hallmarks of the disease and the biodistribution of the administered compound. After screening many compounds in the primary screening assay, potent and selective compounds are identified and tested for efficacy in the animal

model. Lead compounds are then optimized with respect to potency, selectivity, and pharmacokinetic properties. For biologicals, few compounds are involved in the screening assay, but optimization and evaluation criteria are just as stringent as for small molecules. Finally, and most importantly, the compound is evaluated for safety and efficacy in human subjects under those conditions that will apply when marketing the compound as a drug (such as the route of administration, dose level and dosing frequency, and disease population). A substance becomes a drug when it is approved for the marketplace.

Drug discovery, especially for behavioral or affective (central nervous system, CNS) disorders, has flourished since the 1950s. The critical drug discoveries have all followed the paradigm of connecting the disease with the molecular target of pathogenesis and the chemical structure for optimal intervention (illustrated schematically in Figure 12.1).

Traditionally, most affective disorders have been treated with compounds that resemble the neurotransmitters that are deficient or in excess in specific brain regions. The aberrant levels of neurotransmitters (or their receptors), such as norepinephrine, dopamine, acetylcholine, and serotonin, have correlated with behavioral symptoms of schizophrenia, depression, anxiety, sleep disorders, motor dysfunctions, attention difficulties, and cognitive disorders. Most drugs discovered for these disorders resulted from screening compounds directly in rodent behavioral models that mimic the behavior of the disease. In these cases, the "molecular target" or mechanism of action was assumed to be the deficiency or excess of a neurotransmitter.

Later, in the 1960s and 1970s, receptor-binding assays using isolated tissues rich in target receptors became the preferred approach for primary screening but were always followed by evaluation of the active compounds in rodent behavioral models to confirm brain bioavailability and *in vivo* potency of the compounds. Assuming the same desired mechanism, primary screening targets evolved to cell-free preparations of neurotransmitter receptors. Subsequently, in the 1980s and 1990s, complementary DNAs (cDNAs) encoding human receptor subtypes were isolated and expressed to screen for subtype-selective compounds, still using rodent behavioral models for confirmation of efficacy and bioavailability. At the same time, the development of microdialysis techniques with specific neurotransmitter probes allowed the creation of secondary assays that confirmed the desired effect on neurotransmitters and that were more predictive of the behavioral effect and target selectivity.[1–4]

All these methodological advances were used to improve the correlation and specificity of the three elements — chemical structure, molecular target, and disease — leading to the selection of a single preferred compound or drug candidate. The selection of an optimal drug candidate is a key step in the drug discovery process. The appropriate toxicology, formulation studies, and chronic pharmacology studies necessary to start clinical trials are preferably done on the optimal drug candidate. The proper design of clinical trials to ensure approval of the desired claims are not discussed in this chapter, although they are

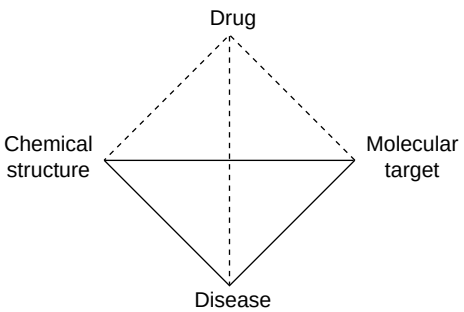


FIGURE 12.1
Schematic representation of drug discovery.

most important for showing efficacy in humans and keeping the cost and time of drug discovery and development to an affordable level.

12.2 Neurodegeneration and Neuroprotective Agents

Since the early 1980s, much effort has focused on animal models of acute and chronic neurodegeneration in search of therapeutics for stroke. Neuronal cell death follows strokes, acute ischemic insults, and chronic neurodegeneration, such as Parkinson's disease, Alzheimer's disease (AD), epilepsy, and Huntington's disease. Up to 80% of all strokes result from focal infarcts and ischemia in the middle cerebral artery (MCA), so the commonly used animal models for neuroprotection are produced by temporary or permanent occlusion of the MCA.[5] Lesions of the MCA include occlusion by electrocoagulation, intraluminal monofilaments, photochemical effects, thrombosis, and endothelin-1, but all of these models necessitate studying reperfusion events and validating MCA occlusion by behavioral assessments.

The major efforts to stop neuronal cell death and provide some neuroprotection have been to block excessive activation of excitatory glutamate receptors or to inhibit apoptosis. Initially, *N*-methyl-D-aspartate (NMDA) antagonists like MK801 were thought to be promising antiischemic agents, but too many side effects and inadequate efficacy in treating cerebral infarcts have prevented approval of these agents. Blockers of ionotropic and metabotropic glutamate receptors have also been clinically inadequate to date. Anti-apoptotic agents may slow the necrosis of neurons, and caspases are key to the induction of neuronal cell death. Of the 16 known caspases, some cause apoptosis by integrating apoptotic signals and activating effector caspases that dismantle neurons. Confirmation of neuroprotection (delayed cell death, not permanent rescue) in animals is still weak, but evidence exists that NMDA, α-amino-3-hydroxy-5-methyl-4-isoxazole propionate acid (AMPA), and metabotropic antagonists can be neuroprotective. A surge of animal studies has been reported recently, and some caspase inhibitors are being evaluated in clinical trials. Long-term treatment with such agents will be required for evidence of neuronal rescue.[6–11]

12.3 Optimizing Chemical Structures

The development of combinatorial chemistry and parallel synthesis allow chemists to synthesize small quantities of relatively pure compounds at rates of a few hundred per week per chemist, in contrast to at least a gram of material for a few compounds per month in previous times. With the parallel development of quantitative structure–activity relationship (SAR) techniques, calculated predictions for oral bioavailability and metabolism, and high-throughput screening (HTS) to directly test compounds in 384-well plates, optimal chemical structures for a primary assay can be achieved in a few months with fewer chemists. These techniques have increased efforts to identify second-generation drugs with improved potency, receptor selectivity, safety, and oral bioavailability, and longer duration of the effects. These techniques have yielded new structural and mechanism classes of antidepressants, antipsychotics, anxiolytics, cognition enhancers, sleep modifiers, and analgesics in the past decade. Additional requirements for CNS-active drugs are moderate lipophilicity,

oral bioavailability, and permeability through the blood-brain barrier. Nasal sprays or dermal patches on the neck can deliver drugs to the brain and reduce first-pass metabolism of drug molecules labile to metabolism (*N*-alkyl groups).

12.4 Target Validation

Even if potent and selective chemical structures exist for a molecular target, the target is not validated for therapeutic intervention until efficacy can be demonstrated in clinical trials. Both the compound and the target are only fully validated for a disease when a drug is registered for the marketplace, which reflects the adequate safety or therapeutic index for a specified patient population by a specific route and frequency of dosing. For example, impairment of cholinergic pathways in the brain has long been believed to result in synaptic disruption and cognitive impairment leading to memory loss, also referred to as the "cholinergic hypothesis." Acetylcholinesterase inhibitors have been approved, such as tacrine (Cognex, Parke-Davis, Ann Arbor); donepezil hydrochloride (Aricept, Eisai/Tokyo and Pfizer/New York); rivastigmine tartrate (Exelon, Novartis Pharmaceuticals, Basel); and galantamine hydrobromide (Reminyl, developed by Johnson & Johnson/New Brunswick and Shire Pharmaceuticals/Wayne, manufactured and marketed by Janssen Pharmaceutica/Titusville), for delaying the progression of cognition or memory loss by allowing significant levels of the neurotransmitter acetylcholine to be sustained in the synapse, but direct cholinergic agonists or mimetics have not yet been approved. Several have been studied clinically; a few have shown efficacy that might validate the target, but none have yet shown a safe therapeutic index to warrant approval. Is it because the drug candidates discovered so far are inadequately selective for brain cholinergic receptors (such as the M1 muscarinic and $\alpha 2\beta 4$ nicotinic receptor subtypes) and cause too many peripheral cholinergic effects? Antimuscarinic drugs have actually been shown to increase Alzheimer's pathology in Parkinson's disease.[12] This is a working hypothesis, but until a more selective compound can be shown to be efficacious and garner approval, stimulation of postsynaptic cholinergic receptors by muscarinic agonists is not a validated therapeutic target.

Glutamate receptors have represented attractive drug discovery targets for two decades because glutamate is the major excitatory neurotransmitter in the brain. Antagonists of glutamate receptors, especially NMDA antagonists, have been studied to reduce conditions like schizophrenia, epilepsy, neuropathic pain, and anxiety, but side effects have halted the development of most of these compounds. Are the compounds not sufficiently selective between the different glutamate receptor subtypes? Do the compounds provoke other activities related to their structure? Are the compounds readily accessible to the critical brain sites? Are the NMDA receptors the critical ones related to these disease conditions? The recent approval of memantine (Namenda, Forest Pharmaceuticals, St. Louis; Axura, Merz$^+$/Germany; and Ebixa, Lundbeck/Copenhagen) for severe cognitive impairment, but not for mild cognitive impairment, does not unequivocally validate the target because memantine is neither a potent nor a selective NMDA modulator, so we cannot conclude that NMDA antagonism is the mechanism underlying its efficacy.

Another problem in validating targets for behavioral disorders related to neurotransmitter abnormalities is the interplay between several neurotransmitter systems in specific brain regions. For example, in the hippocampus, limbic, and nigral–striatal areas, functions connected by serotonin, norepinephrine, and dopamine are interconnected so that blocking selected receptor subtypes or changing synaptic levels of certain neurotransmitters may

affect other transmitters or receptors at the same time. Postsynaptic receptor antagonists may provoke similar behavioral effects as presynaptic agonists or reuptake inhibitors (such as selective serotonin reuptake inhibitors, SSRIs) because presynaptic terminals may be connected to postsynaptic receptors using different neurotransmitters. Furthermore, patients with the same affective disorder may have malfunctions involving different neurotransmitters and require different therapeutic intervention to change the symptoms. All patients with a similar symptom complex may not have the same molecular target malfunction or may not respond well to the same compound or mechanism.

Recently, several groups have clinically evaluated compounds designed to increase or decrease synaptic levels of multiple neurotransmitters concurrently, such as norepinephrine plus dopamine or norepinephrine plus serotonin. The intent is to effectively treat a broader population with similar symptoms. Conversely, Merck and Pfizer's discontinuance of clinical trials of peptide Y antagonists for depression does not disprove the validation of that target until the results are better understood and can answer such questions as: Was efficacy inadequate for a potent and selective compound? Was selectivity of the compound for the target inadequate? Does the molecular target not relate directly to the disease or have too many disconnects in the pathway to the disease?

Validation of a molecular target for a particular disease can shorten the drug discovery process. Discovery groups seek different chemical structures with similar potency and selectivity for a validated target to obtain its own patent protection. They may seek compounds with better brain bioavailability (especially by oral administration), with less-rapid metabolism, or with additional disease claims based on the validated target and mechanism, but all of these drug discovery approaches are of lower risk because the target has been validated. When several drugs are approved in a class, later entries often capture larger market shares if they have fewer side effects or provide better patient compliance or are effective in a broader patient population. The first "blockbuster" in the class requires validation of the target and evidence of safety for the target over long-term dosing and high doses. Expansion of the target to multiple drugs happens quickly after target validation — such as beta blockers, angiotensin-converting enzyme (ACE) inhibitors, H2 antagonists, proton pump inhibitors, 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase inhibitors (statins), and SSRI antidepressants.

Until recently, there have been few targets resulting in blockbusters in CNS and neurologic diseases, primarily because of the lack of validated targets and the reluctance of psychiatrists to prescribe drugs over the long term. The benzodiazepine anxiolytics were the first major class, although their mechanism was not a clearly defined target, and their use decreased when addiction and abuse were associated with the drugs. SSRIs for depression have become the first major target-validated class with blockbuster commercial status because they are relatively safe, not abusable by outpatients, and effective over time without much tolerance build-up. The cholinesterase inhibitors for cognitive impairment have emerged as the next blockbuster class for neuropsychiatric disorders.

12.5 Importance of Biomarkers for Target Validation

The evolution of target validation for blockbuster drugs highlights the need for methods to diagnose specific disease conditions. Most non-CNS categories use quick and simple diagnostic criteria to justify the intervention and follow the course of the disease — blood pressure, heart rate, blood cholesterol level, viral load, and even ulcers by gastroscope. Diagnostic criteria for neuropsychiatric disorders are much more subjective — interviews,

rating scales, impressions, and patient self-diagnosis (pain, anxiety). The lack of objective end points or markers makes it much more difficult to connect a molecular target to a disease condition, especially with some inconsistency between diagnosticians. This results in many physicians "trying" a drug class, but the decision of efficacy and continuing therapy is made largely by the patient. Therefore, compliance is low if patients do not perceive relief from symptoms, if cost or reimbursement becomes a burden, or if they are embarrassed about having the disease. Linking an objective end point or biomarker to the disease status is very important to securing the validation of a molecular target or mechanism. Blood biomarkers have often been inadequate in reflecting pathology of the brain, whereas brain biomarkers require spinal taps for cerebral spinal fluid, and this is not patient-friendly. The recent progress in brain imaging technologies, such as functional magnetic resonance imaging (MRI) and positron emission tomography (PET), offers hope that soon we will be able to measure brain pathology, morphology, and function over time. These imaging techniques may facilitate early diagnosis of patients at risk and enable following the course of disease progression as well as its change in response to an experimental therapeutic. It will be time-consuming and expensive to validate these imaging techniques for the first drug approval, but they will be critical to approval of a claim for reducing pathology.

12.6 Identifying New Molecular Targets

We next explore the process of identifying new molecular targets relevant to neuropsychiatric diseases — whether identification is an attempt to explain the effects of a new compound in a disease or to focus on a particular pathway in the pathophysiology of disease to screen for compounds that optimally interact with the target. As mentioned previously, several drug discovery groups continue to refine selectivity and increase potency for serotonin and catecholamine receptors and uptake transporters, largely because these targets have been clinically validated for depression, schizophrenia, bipolar disorder, anxiety, and attention-deficit disorders. Most of these groups explore novel chemical structures that affect the same mechanisms or molecular targets.

Other neurotransmitter receptors and ion channels have been dissected into subunits or building blocks, and variations of these subunits (even splice variants) have been assembled to make distinct classes of receptor subtypes. The genes for these subunits have been coexpressed and even stably expressed to obtain working quantities of specific receptor subtypes, such as muscarinic, nicotinic (cholinergic ion channels), NMDA, metabotropic, AMPA, kainate, γ-aminobutyrate (GABA), and voltage-gated ion channels. Many research groups have screened these receptor subtypes in transfected cells, even in neuronal cells, against large libraries of drug-like compounds in search of potent and selective ligands. Several compound classes with good activity in animal models have been identified in the past decade, but most have not yet shown clinical efficacy; a few, however, have reached the marketplace. Additional work on these subtype-selective compounds will be carried out during the current decade as we learn the scope and limitations of such compounds for treating patients. Parallel with this drug-screening effort has been an extensive effort to map the genes and receptor densities in the spinal cord and different brain regions to correlate subtype-selective compounds with specific diseases, thus providing a basis for selecting patient populations most likely to respond to these selective compounds.

In this manner, cholinergic activity increased by either circulating levels of acetylcholine through inhibition of acetylcholinesterase (Aricept, Exelon, Reminyl) or by stimulating muscarinic M1 receptors (not yet marketed) has proven beneficial for patients with mild cognitive impairment and early-to-moderate stages of Alzheimer's dementia. Memantine has recently been approved to slow the progression of late-stage AD, so future studies will determine whether it helps patients when the acetylcholinesterase inhibitors cease to work and whether glutamate inhibitors offer an alternative pathway to slow the crippling dementia. A few nicotinic agonists have shown positive effects on cognition but not yet an adequately safe therapeutic index for market approval, again highlighting the "cholinergic hypothesis" for cognition and memory. A few metabotropic subtype blockers have shown clinical activity for anxiety, and an AMPA/kainate blocker has shown good clinical activity in neuropathic pain and migraine[13] as has a voltage-gated calcium channel blocker,[14] but these approaches lack the phase 3 clinical data necessary for target validation, new drug application (NDA) registration, and marketing at this time. Several more of these subtype selective agents working on glutamate, cholinergic, and voltage-gated receptors and ion channels are expected to be validated and approved in the next few years.

At this time, however, we are not aware of any compounds selected primarily by their neuroprotection activity on rodent models that have established clinical efficacy for dementias or related neurodegenerative diseases. This may be partially explained by their priority development for stroke, and clinicians have found it is difficult or unlikely to slow the ischemia in patients if they are not treated aggressively within 3 h of the initial ischemic event. The speed of neurodegeneration in stroke (cerebral ischemia) makes it a much more difficult target for drug intervention than neurodegeneration from slower pathologies such as Alzheimer's, Parkinson's, and malfunctions in neurotransmitters.

A recent pilot study of gene therapy hinted at clinically effective neuroprotection by the nerve growth factor (*NGF*) gene. Skin cells were biopsied from eight mild AD patients and were transfected with a murine leukemia vector that expresses the human *NGF* gene, and subsequently, autologous fibroblasts were injected stereotactically into the nucleus basalis. These patients were followed for 18 to 36 months with no adverse events, and the PET scans and cognitive tests have indicated improved cognitive function. These results suggest that delivering neurotrophic factor genes to critical brain-pathology regions may provide continuing release of growth factors that can sustain neuroprotection and are encouraging for the development of gene therapy as a viable therapeutic approach.[15] By contrast, several growth factors — NGF, brain-derived neurotrophic factor (BDNF), and glial cell-line-derived neurotrophic factor (GDNF) — delivered directly into the brain have not provided clear efficacy.

12.7 Connecting Molecular Targets to Diseases

Returning to the process of drug discovery — which must start with connecting any two of the three discovery elements, most often the compound with the molecular target — we next look briefly at the new techniques and technologies, which may permit the connection of targets to a disease. Chronic and neuropathic pain, but not inflammation, are often associated with excitotoxicity and neurodegeneration. Glutamate antagonists, especially ionotropic glutamate receptors (AMPA/kainate) and NMDA, have been investigated for chronic pain conditions because they block synaptic transmission and because the receptors are located in the dorsal root ganglia. An AMPA/kainate antagonist[13] has shown

clinical efficacy in migraine and lower back pain, but phase 3 trials and a few years of clinical experience are needed before it can be eligible for regulatory approval. It also has shown some neuroprotection in animal studies. Ziconitide (SNX-111, Neurex/Elan, Menlo Park), a snail toxin that blocks the N-type voltage-gated calcium channel, has also shown clinical efficacy for neuropathic pain, blocks synaptic transmission, and is neuroprotective.[14] To link these molecular targets of excitotoxicity with clinical conditions of neuropathic pain, cancer pain, back pain, and migraine, preclinical models of nociception have been appropriate and predictive.[16] The effects of these mechanisms can be seen in rodents with acute pain states provoked by heat, formalin, or electroshock. The effect of the compound on these mechanisms is often more pronounced when the test agent is administered intrathecally. Secondary, *in vivo* tests of hyperalgesia[17] and allodynia[18] have been used to predict the value against neuropathic pain and peripheral nerve injury. Whether glutamate antagonists or calcium channel antagonists provide real alternatives to opiates will depend on their safety, therapeutic index, and patient acceptance, but these neuroprotective agents for chronic pain look promising at this stage of development.

Molecular biologists and geneticists have identified several genes and protein targets that have some correlation to a disease, but a systems approach must be applied for each disease to understand how the different pieces are connected and how each affects the others. Sophisticated genetic techniques have permitted highly penetrant genes to be cloned for early onset forms of Alzheimer's, Parkinson's, and frontotemporal dementias, enhancing our understanding of the etiology of these disorders. Identification of genetic factors with late-onset disorders, such as Alzheimer's, schizophrenia, and bipolar disease, have not yet provided answers to biochemical mechanisms or therapeutic interventions because these genes interact with each other and nongenetic factors to create a much more complex and varied etiology. The systems approach to these late-onset diseases is yielding insights for diagnosis and perhaps prevention as well as therapeutic intervention by focusing on candidate susceptibility genes. The systems approach to late-onset AD will be discussed in the following section to illustrate this approach. Other new technologies that help connect molecular targets to diseases will also be discussed in this textbook, including the use of transgenic mice to approximate disease states caused by aberrant genes and the use of mice in which a gene is knocked out to create a disease model that reflects a missing gene or demonstrates a condition that results from a gene no longer functioning normally.

12.8 Drug Discovery Strategies for Alzheimer's Disease

Alzheimer's Disease is a neurodegenerative disease that manifests as a progressive, cognitive decline with increasing age. Principal pathological hallmarks of AD are senile plaques comprising β-amyloid (Aβ) peptides and neurofibrillary tangles containing hyperphosphorylated microtubule-associated tau protein found in the hippocampus and cerebral cortex of affected individuals. In AD patients, these pathological changes may be evident several years before significant dementia. The 40- and 42-amino-acid variants of Aβ (Aβ_{40} and Aβ_{42}, respectively) are derived from the amyloid precursor protein (APP) by proteolytic cleavage through the sequential activity of β- and γ-secretases (Figure 12.2). The Aβ_{42} variant is highly self-aggregating[19] and normally represents 5 to 10% of the total Aβ secreted from the brain.[20] In early-onset AD, however, an autosomal-dominant form of AD — caused by mutations in APP or presenilin genes (*PS1* and *PS2*) — Aβ_{42} levels are increased several fold, comprising 15 to 40% of the total secreted Aβ.

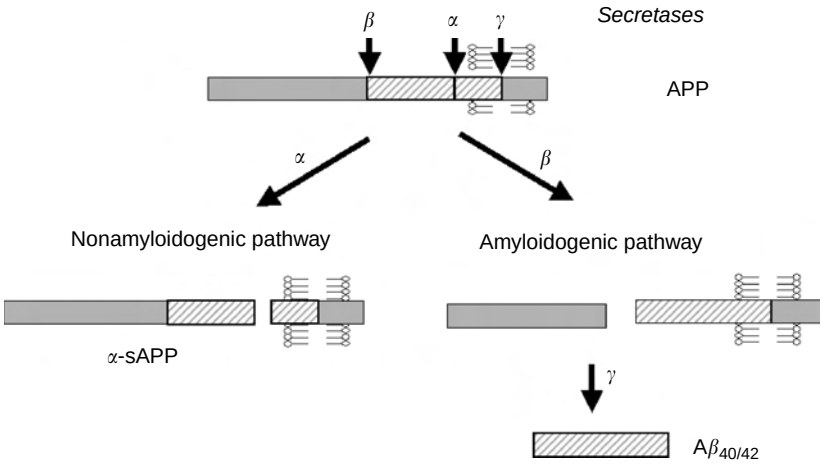


FIGURE 12.2
Generation of Aβ_{42} from APP.

There is cumulative evidence that Aβ, in particular the Aβ_{42} variant, plays a significant role in the pathogenesis of AD, thus forming the basis of the "amyloid hypothesis."[21] This hypothesis predicts that reducing cerebral Aβ levels should prevent or stop the progression of AD and provides a broad framework to formulate a drug discovery strategy, including either inhibition of Aβ production or aggregation, enhancement of its clearance or degradation, or a combination of these approaches.

Considering the amyloid hypothesis as the basis of AD drug discovery, Aβ represents a relevant and measurable end point. Although many questions remain unanswered, recent advances have given us important insights regarding the regulation of Aβ production and degradation, the dynamics of its aggregation, and the cellular mechanism of its neurotoxicity. In the continuum from the initial proteolytic processing of APP to accumulation of plaques, there are a multitude of steps that can regulate the steady-state levels of Aβ in the brain and thus, potentially, represent points of therapeutic intervention. Figure 12.3 schematically depicts this cascade, with each step presenting different opportunities and challenges for drug discovery.

12.8.1 Inhibiting Production

Inhibiting the production of the Aβ peptides represents the most direct approach to curtailing their potential to accumulate as amyloid plaques, by inhibiting either the β-secretase at step 1 or the γ-secretase at step 2. Because these are enzymatic reactions with measurable products, a biochemical assay using a purified enzyme preparation can be integrated into an HTS platform, facilitating the rapid evaluation of large numbers of compounds for inhibition of the enzymatic activity.

With β-secretase, all enzymatic activity can be attributed to BACE1, a membrane-bound aspartyl protease, making this enzyme a valid therapeutic target. Deleting the *BACE1* gene in mice overexpressing human APP leads to marked reduction in cerebral Aβ_{40} and Aβ_{42} and concomitant improvement of Aβ-dependent hippocampal memory deficits.[22,23] Despite extensive compound screening in biochemical assays using recombinant purified BACE1 preparations, pharmacological inhibition of β-secretase *in vivo* has not yet been demonstrated. The lack of success in finding small molecule inhibitors of BACE1 with

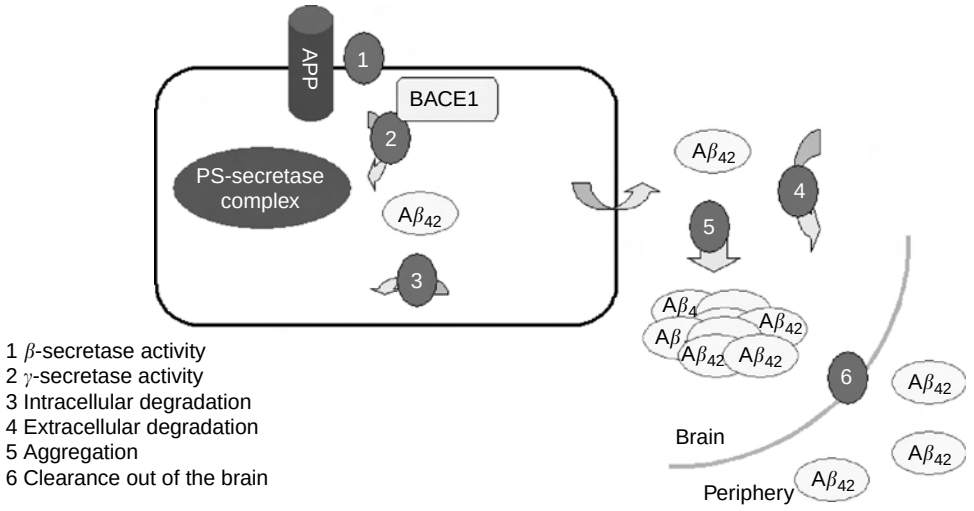


FIGURE 12.3
Multiple mechanisms underlie the accumulation of extracellular $A\beta_{42}$.

in vivo efficacy is partly attributed to the large, active-site pocket of the enzyme that must be accommodated[24] and partly to the limited accessibility of the enzyme that is localized within intracellular vesicles. Nonetheless, inhibition of BACE1 remains a valid therapeutic strategy for AD, and pharmaceutical companies are aggressively pursuing discovery of small molecule BACE1 inhibitors.

With γ-secretase, the enzyme activity resides in a large, multiprotein complex containing at least four identified membrane-bound proteins (presenilin 1, nicastrin, APH-1, and PEN-2). The functional assembly of this complex involves multiple protein–protein interactions, making biochemical HTS strategies less feasible. Furthermore, γ-secretase has multiple substrates, and most "active site" inhibitors of this enzyme are not selective for APP but inhibit the processing of other substrates. One of these substrates is Notch, a critical protein involved in cell differentiation and cell fate. Thus, there is a high level of concern about the potential side effects of γ-secretase inhibitors when used in humans. Recently, it has been reported that certain NSAIDs (sulindac sulfide, ibuprofen, indomethacin) preferentially reduce $A\beta_{42}$ levels relative to $A\beta_{40}$[25–27] without affecting the processing of other γ-secretase substrates. The activity of NSAIDs on γ-secretase does not parallel their antiinflammatory activity but more closely tracks with inhibition of Rho and its effector, Rho-associated kinase (ROCK).[28] Thus, it may be feasible to optimize the activity of the NSAIDs for Rho inhibition, potentially eliminating the gastrointestinal and renal side effects associated with prolonged NSAID usage.

12.8.2 Enhancing Degradation

Enzymatic degradation of Aβ peptides is not as well characterized as the enzymatic formation of these peptides. Recent *in vitro* and *in vivo* data support a functional role for neprilysin and insulin-degrading enzyme (IDE) in this context. As discussed below, genetic data reveal significant linkage and association of IDE with late-onset AD.[29] More recently, analysis of mice expressing an IDE or neprilysin transgene in combination with the human APP transgene revealed that chronically increased expression of either enzyme in neurons of transgenic mice was accompanied by a reduction in brain Aβ,

plaque formation, and related pathology as well as prevention of premature death.[30] These results provide a sound rationale for a drug discovery strategy based on screening compounds for enhancement of the enzymatic activity of either IDE or neprilysin by altering the K_m or V_{max} values.

Activators have been found for other enzymes as part of a drug-screening strategy, for example, activators of glucokinase as potential therapeutic agents for diabetes.[31] Furthermore, Song et al.[32] have reported that certain peptide substrates of IDE, such as dynorphin B, can enhance the proteolysis of Aβ by IDE, without affecting the cleavage of insulin. The authors propose a cooperative mechanism in which the peptide substrate stabilizes the dimeric form of IDE, which has a lower K_m for Aβ. Thus, it is possible to screen compounds for stabilization of the dimer or disruption of the tetramer. Because IDE and neprilysin have other substrates in addition to Aβ, the challenge of selective enhancement of the degradation of Aβ without affecting the proteolysis of other substrates remains. Although the studies in transgenic mice overexpressing either enzyme did not reveal adverse effects during 14 months, the implication of long-term treatment of humans with drugs that activate either peptidase activity remains an open question.

12.8.3 Inhibiting Aggregation

The amyloidogenic Aβ peptide self-associates to form aggregates ranging from soluble oligomers to insoluble amyloid fibrils. Aβ_{42} forms fibrils faster than the Aβ_{40} variant[19] and also appears to be more neurotoxic than Aβ_{40}.[33] It is generally agreed that Aβ aggregation is a toxic event and must be avoided, but it is not clear where the optimal point of therapeutic intervention resides along the aggregation axis. In neuronal cultures, the toxic effects of the Aβ peptide require aggregation, whereas the monomeric forms of the peptide are innocuous.[34,35] Because of the presence of mixed populations of different aggregate forms *in vivo*, it has been difficult to correlate specific neurotoxic effects with any particular oligomeric species. Early studies focused on the insoluble Aβ fibrils as the primary neurotoxic species.[36] However, recent evidence indicates that soluble, oligomeric Aβ species, including dimers and trimers, can inhibit long-term potentiation (LTP) in rats[37] and soluble Aβ oligomers are elevated in the brain of AD patients.[38] Consistent with this, in BACE1 $-/-$ mice, overexpressing human APP, cognitive deficits correlate more closely with the formation of soluble Aβ oligomers than the formation of insoluble amyloid plaques.[23] Furthermore, passive immunization of the PDAPP transgenic mouse (overexpressing human APPV717F) with an anti-Aβ antibody rapidly reversed memory deficits in two behavioral tests, likely by sequestering soluble Aβ species without significantly affecting cerebral amyloid deposits.[39]

The discovery of inhibitors of fibrillogenesis has been hampered by the complexity of the folding pathway, the presence of mixed populations of aggregates, and the lack of sensitive markers of each oligomeric species.[40] Screening assays based on thioflavin-S fluorescence or Congo-red absorbance have been used to identify inhibitors of aggregation of synthetic Aβ peptide. The reported compounds suffer from relatively low potency (effective at <10 μM) and "flat" SARs.[41–43]

A viable drug discovery strategy can be formulated if a rapid screening assay can be established to monitor the formation of specific Aβ oligomers under more physiological conditions. Wigley et al.[44] have described an assay that relies on structural complementation between the α- and ω-fragments of β-galactosidase. The functional β-galactosidase is generated when the two fragments come together in a soluble state, enzyme activity being monitored readily by color change. The α-fragment is fused to the Aβ peptide, and the aggregation of Aβ results in inactive β-galactosidase. Thus, aggregation inhibitors can be identified by the restoration of β-galactosidase activity using a facile, colorimetric assay.

Another approach to inhibition of aggregation involves metal chelators, based on the observations that the binding of copper to Aβ promotes self-aggregation of the peptide.[45] One such compound is clioquinol that promotes solubilization and clearance of Aβ by binding zinc and copper ions and thus inhibiting their binding to Aβ. Results from a pilot phase 2 clinical trial suggest that clioquinol improves cognition and lowers plasma levels of Aβ_{42} in some patients.

12.8.4 Enhancing Clearance

Clearing the Aβ peptide from the brain has also received attention as a therapeutic approach. Active immunization with Aβ peptide was effective in stimulating Aβ clearance *in vivo*, as evidenced by the disappearance of amyloid plaques in the PDAPP transgenic mouse model of AD.[46,47] Furthermore, there were quantifiable improvements in behavioral tests of cognition in mice.[48] Active immunization was also effective in slowing down the progression of dementia in AD patients, although clinical trials were stopped because of the occurrence of aseptic meningoencephelitis in 6% of the patients.[49,50] Passive immunization of transgenic mice using a 5-month treatment with anti-Aβ antibodies also prevented the formation of amyloid plaques in the CNS[47,51] and in other studies,[39] resulted in unexpected reversal of memory deficits in certain learning and memory tests in the PDAPP mouse within 24 h of injection. There was a concomitant increase in Aβ_{40}, Aβ_{42}, and Aβ-antibody complexes in the CSF and plasma. These results suggest that antibody-mediated sequestration of soluble Aβ aggregates can lead to rapid improvement in memory, likely by facilitating the efflux of Aβ from the brain down its concentration gradient, referred to as the "peripheral sink effect." Because these antibodies do not permeate into the brain, the passive immunization strategy avoids CNS complications that are encountered in active immunization.

12.8.5 Animal Models

The AD animal models primarily involve transgenic mice that overexpress wild-type APP or one of the familial Alzheimer's disease (FAD)-mutant variants of human APP, in some cases in combination with a FAD-mutant presenilin.[48,52,53] In all these cases, the transgenic animal exhibits high levels of cerebral and plasma Aβ peptides and age-dependent formation of cerebral Aβ amyloid deposits that also contain neurites, astrocytes, and microglia. The severity of cerebral amyloidosis correlates with the level of transgene expression and the particular mutant form of APP or PS expressed. Although these models do not display neurofibrillary tangles, the triple transgenic model harboring mutant PS, mutant APP, and mutant tau[54] does exhibit both amyloid plaques and tangles. The different genetically engineered, AD mouse strains display cognitive deficits in specific rodent models of learning and memory. In particular, cholinergic regulation of hippocampal neurons is impaired in the APP-transgenic mice (Tg2576 mouse overexpressing human APP695$_{sw}$), but these cholinergic deficits can be rescued by deletion of BACE1.[23] Temporal dissection of the biochemical, physiological, pathological, and behavioral end points has indicated that the synaptic dysfunction precedes Aβ deposition which, in turn, precedes tau pathology.[54] This points to soluble Aβ as the critical factor affecting alterations in synaptic function before the deposition of insoluble Aβ aggregates.

12.8.6 Cholesterol and Alzheimer's Disease

There is increasing evidence based on genetics, epidemiology, and more recently, cell biology that implicates cholesterol in the pathogenesis of AD.[55,56] The principal cholesterol-carrier

protein in the brain is Apolipoprotein E (ApoE), and the Apoε4 allele is the only confirmed genetic risk factor for late-onset AD. ApoE and ApoJ (also called clusterin), two abundantly expressed apolipoproteins in the brain, both act as Aβ chaperones and cooperatively suppress Aβ levels and deposition in the brain.[57] In addition to its role as a chaperone, ApoE also regulates the metabolism of Aβ in the brain extracellular space.[57] There are increasing genetic data (see the next section) that implicate several other genes in the cholesterol pathway, such as cholesterol 24-hydroxylase, in AD. In the cell, cholesterol impacts membrane structure, and APP, BACE1, and all four components of the γ-secretase complex (presenilin 1, nicastrin, APH-1, and PEN-2) are membrane-associated proteins. Reducing cholesterol esters by inhibiting acyl-coenzyme A cholesterol acyltransferase (ACAT) reduces Aβ secretion.[58] Accumulation of intracellular cholesterol is accompanied by accumulation of presenilin 1 in the endosomal pathway and increased Aβ production.[59] Despite significant correlative data, the molecular mechanism of how cholesterol affects Aβ levels is not yet fully understood. The epidemiological evidence comes from the observations that the incidence of AD is reduced by almost 70% among statin users.[56,60] Definitive clinical efficacy trials are currently underway with several statins, such as simvastatin (phase 3) and atorvastatin (phase 2).

12.8.7 Genetics of Alzheimer's Disease

Ultimately, identification of specific proteins involved in the pathogenesis of late-onset AD is required to discover selective and efficacious therapeutics. The characterization of disease-causing mutations in APP and presenilin genes associated with the autosomal-dominant early-onset familial AD led to the identification of the first generation of molecular targets in AD biology (see the review by Hardy[61]). The genetics of late-onset AD is more complex and heterogeneous, and the discovery of molecular targets in this more common form of the disease has been more challenging. ApoE is the only established genetic risk factor for late-onset AD, the ε4 allele being more prevalent in AD patients when compared with cognitively healthy, age-matched control subjects.[62] The results with ApoE-ε4 have been widely replicated, and there is consensus that this allele acts to decrease the age of AD onset. Daw et al.[63] modeled AD as a quantitative trait locus using age of onset, and they predicted as many as seven additional late-onset AD genes, with one of these having greater effect on the disease than ApoE. Several family-based genetic association studies have been undertaken with late-onset AD, yielding potential molecular targets on several chromosomes (see reviews 64 and 65). Among these is IDE (discussed earlier) located on chromosome 10q23, a region with significant linkage to late-onset AD.[29] Four genome-wide screens also detected linkage in this region of chromosome 10,[66–69] supporting the existence of a gene involved in late-onset AD at this location. More recent advances (announced at the 33rd Annual Meeting of the Society for Neuroscience) revealed other candidate genes, such as α-T catenin, glutathione S-transferase, omega-1, and cholesterol-25-hydroxylase, all on chromosome 10; ubiquilin on chromosome 9; and cholesterol 24-hydroxylase on chromosome 14.

Biological validation of these and other genes yet to be discovered is essential for formulating a sound drug discovery strategy. Several converging technologies are available to facilitate the biological validation of candidate genes in cells (recombinant expression and RNAi knock out, for instance) and whole animals (such as regulated transgene expression and gene knock out). Pursuant to confirmation of the biological role of the candidate gene in AD pathways, a more mechanistic and target-based drug discovery strategy can be implemented. Drugs based on a specific molecular mechanism may have fewer side effects, a much-desired outcome for prolonged treatment of a neurodegenerative disease such as AD.

References

1. Reid, R.T., Lloyd, G.K., and Rao, T.S., Pharmacological characterization of nicotine-induced acetylcholine release in the rat hippocampus in vivo: evidence for a permissive dopamine synapse, *Br. J. Pharmacol.*, 127, 1486, 1999.
2. Day, J.C. and Fibiger, H.C., Dopaminergic regulation of septohippocampal cholinergic neurons, *J. Neurochem.*, 63, 2086, 1994.
3. Imperato, A., Mulas, A., and Di Chiara, G., Nicotine preferentially stimulates dopamine release in the limbic system of freely moving rats, *Eur. J. Pharmacol.*, 132, 337, 1986.
4. Moghaddam, B. and Bunney, B.S., Ionic composition of microdialysis perfusing solution alters the pharmacological responsiveness and basal outflow of striatal dopamine, *J. Neurochem.*, 53, 652, 1989.
5. O'Neill, M., Rodent models of focal cerebral ischemia, in *Current Protocols in Neuroscience*, Crawley, J., Gerfen C., Rogawski, M., Sibley, D., Skolnick, P. and Wray, S., Eds., Wiley, New York, 1998, chap. 9.
6. Budd, S. and Rao, T., Excitotoxicity in acute and chronic neurodegeneration, *Curr. Opin. CPNS Investig. Drugs*, 2, 403, 2000.
7. Lee, J.M., Zipfel, G.J., and Choi, D.W., The changing landscape of ischaemic brain injury mechanisms, *Nature*, 399, A7, 1999.
8. Monaghan, D.T., Bridges, R.J., and Cotman, C.W., The excitatory amino acid receptors: their classes, pharmacology, and distinct properties in the function of the central nervous system, *Annu. Rev. Pharmacol. Toxicol.*, 29, 365, 1989.
9. Endres, M., Namura, S., Shimizu-Sasamata, M., Waeber, C., Zhang, L., Gomez-Isla, T., Hyman, B.T., and Moskowitz, M.A., Attenuation of delayed neuronal death after mild focal ischemia in mice by inhibition of the caspase family, *J. Cereb. Blood Flow Metab.*, 18, 238, 1998.
10. Ona, V.O., Li, M., Vonsattel, J.P., Andrews, L.J., Khan, S.Q., Chung, W.M., Frey, A.S., Menon, A.S., Li, X.J., Stieg, P.E., Yuan, J., Penney, J.B., Young, A.B., Cha, J.H., and Friedlander, R.M., Inhibition of caspase-1 slows disease progression in a mouse model of Huntington's disease, *Nature*, 399, 263, 1999.
11. Namura, S., Zhu, J., Fink, K., Endres, M., Srinivasan, A., Tomaselli, K.J., Yuan, J., and Moskowitz, M.A., Activation and cleavage of caspase-3 in apoptosis induced by experimental cerebral ischemia, *J. Neurosci.*, 18, 3659, 1998.
12. Perry, E.K., Kilford, L., Lees, A.J., Burn, D.J., and Perry, R.H., Increased Alzheimer pathology in Parkinson's disease related to antimuscarinic drugs, *Ann. Neurol.*, 54, 235, 2003.
13. Gilron, I., Max, M.B., Lee, G., Booher, S.L., Sang, C.N., Chappell, A.S., and Dionne, R.A., Effects of the 2-amino-3-hydroxy-5-methyl-4-isoxazole-proprionic acid/kainate antagonist LY293558 on spontaneous and evoked postoperative pain, *Clin. Pharmacol. Ther.*, 68, 320, 2000.
14. Staats, P.S., Yearwood, T., Charapata, S.G., Presley, R.W., Wallace, M.S., Byas-Smith, M., Fisher, R., Bryce, D.A., Mangieri, E.A., Luther, R.R., Mayo, M., McGuire, D., and Ellis, D., Intrathecal ziconotide in the treatment of refractory pain in patients with cancer or AIDS: a randomized controlled trial, *JAMA*, 291, 63, 2004.
15. Tuszynski, M.H. and Blesch, A., Nerve growth factor: from animal models of cholinergic neuronal degeneration to gene therapy in Alzheimer's disease, *Prog. Brain. Res.*, 146, 441, 2004.
16. Yaksh, T., Preclinical models of nociception, in *Anesthesia: Biological Foundations*, Yaksh, T., Lynch, C., Zapol, W., Maze, M., Biebuyck, J.F. and Saidman, L.J., Eds., Lippincott-Raven Publishers, Philadelphia, 1997, chap. 40.
17. Bennett, G.J. and Xie, Y.K., A peripheral mononeuropathy in rat that produces disorders of pain sensation like those seen in man, *Pain*, 33, 87–107, 1988.
18. Choi, Y., Yoon, Y.W., Na, H.S., Kim, S.H., and Chung, J.M., Behavioral signs of ongoing pain and cold allodynia in a rat model of neuropathic pain, *Pain*, 59, 360–376, 1994.
19. Jarrett, J.T., Berger, E.P., and Lansbury, P.T., Jr., The C-terminus of the beta protein is critical in amyloidogenesis, *Ann. NY Acad. Sci.*, 695, 144, 1993.

20. Bitan, G., Kirkitadze, M.D., Lomakin, A., Vollers, S.S., Benedek, G.B., and Teplow, D.B., Amyloid beta-protein (Abeta) assembly: Abeta 40 and Abeta 42 oligomerize through distinct pathways, *Proc. Natl. Acad. Sci. USA,* 100, 330, 2003.
21. Hardy, J. and Selkoe, D.J., The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics, *Science*, 297, 353, 2002.
22. Luo, Y., Bolon, B., Kahn, S., Bennett, B.D., Babu-Khan, S., Denis, P., Fan, W., Kha, H., Zhang, J., Gong, Y., Martin, L., Louis, J.C., Yan, Q., Richards, W.G., Citron, M., and Vassar, R., Mice deficient in BACE1, the Alzheimer's beta-secretase, have normal phenotype and abolished beta-amyloid generation, *Nat. Neurosci*., 4, 231, 2001.
23. Ohno, M., Sametsky, E.A., Younkin, L.H., Oakley, H., Younkin, S.G., Citron, M., Vassar, R., and Disterhoft, J.F., BACE1 deficiency rescues memory deficits and cholinergic dysfunction in a mouse model of Alzheimer's disease, *Neuron*, 41, 27, 2004.
24. John, V., Beck, J.P., Bienkowski, M.J., Sinha, S., and Heinrikson, R.L., Human beta-secretase (BACE) and BACE inhibitors, *J. Med. Chem*., 46, 4625, 2003.
25. Weggen, S., Eriksen, J.L., Das, P., Sagi, S.A., Wang, R., Pietrzik, C.U., Findlay, K.A., Smith, T.E., Murphy, M.P., Bulter, T., Kang, D.E., Marquez-Sterling, N., Golde, T.E., and Koo, E.H., A subset of NSAIDs lower amyloidogenic Abeta42 independently of cyclooxygenase activity, *Nature*, 414, 212, 2001.
26. Weggen, S., Eriksen, J.L., Sagi, S.A., Pietrzik, C.U., Golde, T.E., and Koo, E.H., Abeta42-lowering nonsteroidal anti-inflammatory drugs preserve intramembrane cleavage of the amyloid precursor protein (APP) and ErbB-4 receptor and signaling through the APP intracellular domain, *J. Biol. Chem*., 278, 30748, 2003.
27. Weggen, S., Eriksen, J.L., Sagi, S.A., Pietrzik, C.U., Ozols, V., Fauq, A., Golde, T.E., and Koo, E.H., Evidence that nonsteroidal anti-inflammatory drugs decrease amyloid beta 42 production by direct modulation of gamma-secretase activity, *J. Biol. Chem*., 278, 31831, 2003.
28. Zhou, Y., Su, Y., Li, B., Liu, F., Ryder, J.W., Wu, X., Gonzalez-DeWhitt, P.A., Gelfanova, V., Hale, J.E., May, P.C., Paul, S.M., and Ni, B., Nonsteroidal anti-inflammatory drugs can lower amyloidogenic Abeta42 by inhibiting Rho, *Science*, 302, 1215, 2003.
29. Bertram, L., Blacker, D., Mullin, K., Keeney, D., Jones, J., Basu, S., Yhu, S., McInnis, M.G., Go, R.C., Vekrellis, K., Selkoe, D.J., Saunders, A.J., and Tanzi, R.E., Evidence for genetic linkage of Alzheimer's disease to chromosome 10q, *Science*, 290, 2302, 2000.
30. Leissring, M.A., Farris, W., Chang, A.Y., Walsh, D.M., Wu, X., Sun, X., Frosch, M.P., and Selkoe, D.J., Enhanced proteolysis of beta-amyloid in APP transgenic mice prevents plaque formation, secondary pathology, and premature death, *Neuron*, 40, 1087, 2003.
31. Grimsby, J., Sarabu, R., Corbett, W.L., Haynes, N.E., Bizzarro, F.T., Coffey, J.W., Guertin, K.R., Hilliard, D.W., Kester, R.F., Mahaney, P.E., Marcus, L., Qi, L., Spence, C.L., Tengi, J., Magnuson, M.A., Chu, C.A., Dvorozniak, M.T., Matschinsky, F.M., and Grippo, J.F., Allosteric activators of glucokinase: potential role in diabetes therapy, *Science*, 301, 370, 2003.
32. Song, E.S., Juliano, M.A., Juliano, L., and Hersh, L.B., Substrate activation of insulin-degrading enzyme (insulysin). A potential target for drug development, *J. Biol. Chem*., 278, 49789, 2003.
33. Dahlgren, K.N., Manelli, A.M., Stine, W.B., Jr., Baker, L.K., Krafft, G.A., and LaDu, M.J., Oligomeric and fibrillar species of amyloid-beta peptides differentially affect neuronal viability, *J. Biol. Chem*., 277, 32046, 2002.
34. Pike, C.J., Burdick, D., Walencewicz, A.J., Glabe, C.G., and Cotman, C.W., Neurodegeneration induced by beta-amyloid peptides in vitro: the role of peptide assembly state, *J. Neurosci*., 13, 1676, 1993.
35. Lorenzo, A. and Yankner, B.A., Beta-amyloid neurotoxicity requires fibril formation and is inhibited by congo red, *Proc. Natl. Acad. Sci. USA*, 91, 12243, 1994.
36. Hardy, J.A. and Higgins, G.A., Alzheimer's disease: the amyloid cascade hypothesis, *Science*, 256, 184, 1992.
37. Walsh, D.M., Klyubin, I., Fadeeva, J.V., Cullen, W.K., Anwyl, R., Wolfe, M.S., Rowan, M.J., and Selkoe, D.J., Naturally secreted oligomers of amyloid beta protein potently inhibit hippocampal long-term potentiation in vivo, *Nature*, 416, 535, 2002.
38. Gong, Y., Chang, L., Viola, K.L., Lacor, P.N., Lambert, M.P., Finch, C.E., Krafft, G.A., and Klein, W.L., Alzheimer's disease-affected brain: presence of oligomeric A beta ligands (ADDLs)

suggests a molecular basis for reversible memory loss, *Proc. Natl. Acad. Sci. USA*, 100, 10417, 2003.
39. Dodart, J.C., Bales, K.R., Gannon, K.S., Greene, S.J., DeMattos, R.B., Mathis, C., DeLong, C.A., Wu, S., Wu, X., Holtzman, D.M., and Paul, S.M., Immunization reverses memory deficits without reducing brain Abeta burden in Alzheimer's disease model, *Nat. Neurosci*., 5, 452, 2002.
40. LeVine, H., The challenge of inhibiting Abeta polymerization, Curr. Med. Chem., 9, 1121, 2002.
41. Tomiyama, T., Shoji, A., Kataoka, K., Suwa, Y., Asano, S., Kaneko, H., and Endo, N., Inhibition of amyloid beta protein aggregation and neurotoxicity by rifampicin: its possible function as a hydroxyl radical scavenger, *J. Biol. Chem*., 271, 6839, 1996.
42. Wood, S.J., MacKenzie, L., Maleeff, B., Hurle, M.R., and Wetzel, R., Selective inhibition of Abeta fibril formation, *J. Biol. Chem*., 271, 4086, 1996.
43. Kuner, P., Bohrmann, B., Tjernberg, L.O., Naslund, J., Huber, G., Celenk, S., Gruninger-Leitch, F., Richards, J.G., Jakob-Roetne, R., Kemp, J.A., and Nordstedt, C., Controlling polymerization of beta-amyloid and prion-derived peptides with synthetic small molecule ligands, *J. Biol. Chem*., 275, 1673, 2000.
44. Wigley, W.C., Stidham, R.D., Smith, N.M., Hunt, J.F., and Thomas, P.J., Protein solubility and folding monitored in vivo by structural complementation of a genetic marker protein, *Nat. Biotechnol*., 19, 131, 2001.
45. Bush, A.I., Masters, C.L., and Tanzi, R.E., Copper, beta-amyloid, and Alzheimer's disease: tapping a sensitive connection, *Proc. Natl. Acad. Sci. USA*, 100, 11193, 2003.
46. Schenk, D., Barbour, R., Dunn, W., Gordon, G., Grajeda, H., Guido, T., Hu, K., Huang, J., Johnson-Wood, K., Khan, K., Kholodenko, D., Lee, M., Liao, Z., Lieberburg, I., Motter, R., Mutter, L., Soriano, F., Shopp, G., Vasquez, N., Vandevert, C., Walker, S., Wogulis, M., Yednock, T., Games, D., and Seubert, P., Immunization with amyloid-beta attenuates Alzheimer-disease-like pathology in the PDAPP mouse, *Nature*, 400, 173, 1999.
47. Bard, F., Cannon, C., Barbour, R., Burke, R.L., Games, D., Grajeda, H., Guido, T., Hu, K., Huang, J., Johnson-Wood, K., Khan, K., Kholodenko, D., Lee, M., Lieberburg, I., Motter, R., Nguyen, M., Soriano, F., Vasquez, N., Weiss, K., Welch, B., Seubert, P., Schenk, D., and Yednock, T., Peripherally administered antibodies against amyloid beta-peptide enter the central nervous system and reduce pathology in a mouse model of Alzheimer disease, *Nat. Med*., 6, 916, 2000.
48. Janus, C., Chishti, M.A., and Westaway, D., Transgenic mouse models of Alzheimer's disease, *Biochim. Biophys. Acta*, 1502, 63, 2000.
49. Schenk, D., Amyloid-beta immunotherapy for Alzheimer's disease: the end of the beginning, *Nat. Rev. Neurosci*., 3, 824, 2002.
50. Orgogozo, J.M., Gilman, S., Dartigues, J.F., Laurent, B., Puel, M., Kirby, L.C., Jouanny, P., Dubois, B., Eisner, L., Flitman, S., Michel, B.F., Boada, M., Frank, A., and Hock, C., Subacute meningoencephalitis in a subset of patients with AD after Abeta42 immunization, *Neurology*, 61, 46, 2003.
51. DeMattos, R.B., Bales, K.R., Cummins, D.J., Dodart, J.C., Paul, S.M., and Holtzman, D.M., Peripheral anti-A beta antibody alters CNS and plasma A beta clearance and decreases brain A beta burden in a mouse model of Alzheimer's disease, *Proc. Natl. Acad. Sci. USA*, 98, 8850, 2001.
52. Ashe, K.H., Learning and memory in transgenic mice modeling Alzheimer's disease, *Learn Mem*., 8, 301, 2001.
53. Wong, P.C., Cai, H., Borchelt, D.R., and Price, D.L., Genetically engineered mouse models of neurodegenerative diseases, *Nat. Neurosci*., 5, 633, 2002.
54. Oddo, S., Caccamo, A., Shepherd, J.D., Murphy, M.P., Golde, T.E., Kayed, R., Metherate, R., Mattson, M.P., Akbari, Y., and LaFerla, F.M., Triple-transgenic model of Alzheimer's disease with plaques and tangles: intracellular Abeta and synaptic dysfunction, *Neuron*, 39, 409, 2003.
55. Puglielli, L., Tanzi, R.E., and Kovacs, D.M., Alzheimer's disease: the cholesterol connection, *Nat. Neurosci*., 6, 345, 2003.
56. Wolozin, B., Cholesterol and the biology of Alzheimer's disease, *Neuron*, 41, 7, 2004.

57. DeMattos, R.B., Cirrito, J.R., Parsadanian, M., May, P.C., O'Dell, M.A., Taylor, J.W., Harmony, J.A., Aronow, B.J., Bales, K.R., Paul, S.M., and Holtzman, D.M., ApoE and clusterin cooperatively suppress abeta levels and deposition: evidence that ApoE regulates extracellular abeta metabolism in vivo, *Neuron*, 41, 193, 2004.
58. Puglielli, L., Konopka, G., Pack-Chung, E., Ingano, L.A., Berezovska, O., Hyman, B.T., Chang, T.Y., Tanzi, R.E., and Kovacs, D.M., Acyl-coenzyme A: cholesterol acyltransferase modulates the generation of the amyloid beta-peptide, *Nat. Cell Biol.*, 3, 905, 2001.
59. Burns, M., Gaynor, K., Olm, V., Mercken, M., LaFrancois, J., Wang, L., Mathews, P.M., Noble, W., Matsuoka, Y., and Duff, K., Presenilin redistribution associated with aberrant cholesterol transport enhances beta-amyloid production in vivo, *J. Neurosci.*, 23, 5645, 2003.
60. Jick, H., Zornberg, G.L., Jick, S.S., Seshadri, S., and Drachman, D.A., Statins and the risk of dementia, *Lancet*, 356, 1627, 2000.
61. Hardy, J., Amyloid, the presenilins and Alzheimer's disease, *Trends Neurosci.*, 20, 154, 1997.
62. Strittmatter, W.J., Saunders, A.M., Schmechel, D., Pericak-Vance, M., Enghild, J., Salvesen, G.S., and Roses, A.D., Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease, *Proc. Natl. Acad. Sci. USA*, 90, 1977, 1993.
63. Warwick Daw, E., Payami, H., Nemens, E.J., Nochlin, D., Bird, T.D., Schellenberg, G.D., and Wijsman, E.M., The number of trait loci in late-onset Alzheimer disease, *Am. J. Hum. Genet.*, 66, 196, 2000.
64. Bertram, L. and Tanzi, R.E., Dancing in the dark? The status of late-onset Alzheimer's disease genetics, *J. Mol. Neurosci.*, 17, 127, 2001.
65. Tanzi, R.E. and Bertram, L. New frontiers in Alzheimer's disease genetics, *Neuron*, 32, 181, 2001.
66. Ertekin-Taner, N., Graff-Radford, N., Younkin, L.H., Eckman, C., Baker, M., Adamson, J., Ronald, J., Blangero, J., Hutton, M., and Younkin, S.G., Linkage of plasma Abeta42 to a quantitative locus on chromosome 10 in late-onset Alzheimer's disease pedigrees, *Science*, 290, 2303, 2000.
67. Myers, A., Holmans, P., Marshall, H., Kwon, J., Meyer, D., Ramic, D., Shears, S., Booth, J., DeVrieze, F.W., Crook, R., Hamshere, M., Abraham, R., Tunstall, N., Rice, F., Carty, S., Lillystone, S., Kehoe, P., Rudrasingham, V., Jones, L., Lovestone, S., Perez-Tur, J., Williams, J., Owen, M.J., Hardy, J., and Goate, A.M., Susceptibility locus for Alzheimer's disease on chromosome 10, *Science*, 290, 2304, 2000.
68. Li, Y.J., Scott, W.K., Hedges, D.J., Zhang, F., Gaskell, P.C., Nance, M.A., Watts, R.L., Hubble, J.P., Koller, W.C., Pahwa, R., Stern, M.B., Hiner, B.C., Jankovic, J., Allen, F.A., Jr., Goetz, C.G., Mastaglia, F., Stajich, J.M., Gibson, R.A., Middleton, L.T., Saunders, A.M., Scott, B.L., Small, G.W., Nicodemus, K.K., Reed, A.D., Schmechel, D.E., Welsh-Bohmer, K.A., Conneally, P.M., Roses, A.D., Gilbert, J.R., Vance, J.M., Haines, J.L., and Pericak-Vance, M.A., Age at onset in two common neurodegenerative diseases is genetically controlled, *Am. J. Hum Genet.*, 70, 985, 2002.
69. Blacker, D., Bertram, L., Saunders, A.J., Moscarillo, T.J., Albert, M.S., Wiener, H., Perry, R.T., Collins, J.S., Harrell, L.E., Go, R.C., Mahoney, A., Beaty, T., Fallin, M.D., Avramopoulos, D., Chase, G.A., Folstein, M.F., McInnis, M.G., Bassett, S.S., Doheny, K.J., Pugh, E.W., and Tanzi, R.E., Results of a high-resolution genome screen of 437 Alzheimer's disease families, *Hum. Mol. Genet.*, 12, 23, 2003.

13

Safety Pharmacology: Past, Present, and Future

Jean-Pierre Valentin and Tim G. Hammond

CONTENTS

13.1 Background

13.1.1 Reasons for Drug Failure in Clinical Development/Withdrawal from Market Places

The reasons for drug attrition have evolved over the years; over the last decade, lack of safety (both non-clinical and clinical) remains the major cause of attrition during clinical development, which accounts for approximately 35 to 40% of all drug discontinuation (see Table 13.1).[1a–1c] More worrying is the fact that there is no clear trend toward a reduction of the attrition owing to safety reasons.

In this section, a brief summary of the nature, frequency, and consequences of adverse drug reactions (ADRs) in two clinical situations is presented. There are ADRs experienced by healthy volunteers and patients participating in clinical studies with potential new medicines and those experienced by patients who are prescribed licensed medicines. A review of these two situations points to areas of success with the current practices for non-clinical safety pharmacology testing but also identifies some areas where further research might lead to new or better safety pharmacology tests. Prior to reviewing the literature, some

TABLE 13.1

Evolution of Reasons for Termination of Drug Development

Reasons	Percentage (Kennedy, 1997)[1c]	Percentage (Lasser et al., 2002)[1b]
Portfolio considerations	23	22
Clinical efficacy	22	23
Clinical safety	12	20
Toxicology[a]	23	19
Other[b]	20	16

[a] Includes general/safety pharmacology.

[b] Includes reasons such as clinical pharmacokinetics/bioavailability, non-clinical efficacy, non-clinical pharmacokinetics/bioavailability, formulation, patent, legal or commercial, and regulatory. Overall, safety reasons accounted for up to ~40% of all discontinuation of drug development.

Source: Adapted from Lasser, K.E. et al., *JAMA*, 287, 2215–2220, 2002 and Kennedy, T., *Drug Discov. Dev.*, 2, 436–444, 1997.

definitions are worth considering. An *adverse event* (AE) is defined as an unintended injury caused by medical management rather than the disease process. ADRs are a subset of AEs, which are thought to be causally related to the use of a medicine. A serious AE or ADR results in death, is life threatening, requires hospitalization, results in persistent disability, and is a congenital abnormality. The severity can be classified as (1) mild — slightly bothersome; (2) relieved with symptomatic treatment; (3) moderate — bothersome, interferes with activities, only partially relieved with symptomatic treatment; or (4) severe — prevents regular activities, not relieved with symptomatic treatment. So, a serious ADR is always significant and has a high impact — it can lead to the discontinuation of a drug in development, a significant limitation in the use of a drug (precaution, contra-indication), or even to the withdrawal of the drug from the market place. A nonserious ADR can be more or less severe in its intensity and its impact will depend upon its frequency and intensity. The impact of serious and nonserious ADRs on a drug's commercial success will be titrated against the overall clinical benefit the drug brings to the patient.

ADRs in humans fall into five types. Of these, acute safety pharmacology studies can reasonably be expected to predict Type A ADRs (see Table 13.2).[2] This means that ~75% of clinical ADRs are potentially predictable on the basis of non-clinical safety pharmacology studies.

During early development, the first clinical studies (Phases I and IIA) are generally very safe.[2,3] In fact, molecules with a significant potential to generate serious ADRs are probably never given to healthy volunteers and given to patients (e.g., refractory cancer patients) only with great care. These studies are to be conducted diligently with careful monitoring for the emergence of potentially worrisome ADRs. While adverse events do occur, they are generally more related to the experimental procedures (e.g., needle puncture) than to the drugs. Safety pharmacology probably contributes significantly to the maintenance of this good track record. This is supported by published reports showing that single-dose non-clinical safety studies could overall accurately predict the clinical outcome.[4–6] The common ADRs observed with a high incidence (10–30%) during these phases are linked to the gastrointestinal and central nervous systems. In addition, ADRs that occurred with a low incidence are also detected. They are often specific to the new chemical entity (NCE) under investigation and are often pharmacologically mediated.

TABLE 13.2

Classification of ADRs in Humans

Type A	Dose-dependent; predictable from primary, secondary, and safety pharmacology	Main cause of ADRs (~75%), rarely lethal
Type B	Idiosyncratic response, not predictable, not dose-related	Responsible for ~25% of ADRs, but majority of lethal ones
Type C	Long-term adaptive changes	Commonly occurs with some classes of drug
Type D	Delayed effects (e.g., carcinogenicity, teratogenicity)	Low incidence
Type E	Rebound effects following discontinuation of therapy	Commonly occurs with some classes of drug

Note: Conventional safety pharmacology studies can reasonably be expected to predict "Type A" ADRs. Functional toxicological measurements may predict "Type C" ADRs. Conventional toxicology studies address "Type D" ADRs. Prediction of "Type B" responses requires a more extensive non-clinical and clinical evaluation, often only addressing risk factors for the idiosyncratic response. "Type E" ADRs are rarely investigated non-clinically using functional measurements unless there is cause for concern.

Source: Adapted from Redfern, W.S. et al., *Fundam. Clin. Pharmacol.*, 16, 161–173, 2002.

TABLE 13.3
Major Causes of Acute Functional Adverse Drug Reactions

Acute Adverse Drug Reaction	Example
Augmented ("supratherapeutic") effect of interaction with the primary molecular target	Pronounced bradycardia with a beta blocker; pronounced hypotension with an angiotensin II receptor antagonist
Interaction with the primary molecular target present in nontarget tissues	Sedation caused by antihistamines
Interactions with secondary molecular targets	Interactions with the hERG cardiac channel leading to QT interval prolongation (e.g., some antipsychotics and antihistamines drugs)
Nonspecifics effects	
Pharmacologically active metabolites	

During Phase II/III clinical development, a large number (often the majority) of patients report adverse events, with a wide variation in the type, frequency, and severity of events detected. Nonserious ADRs are often mechanism-, or drug class-, or disease-related (Table 13.3). Such ADRs limit the utility of a new medicine by restricting its use to those patients who either do not experience or can tolerate the ADRs, and they do not usually pose a safety issue. Serious ADRs tend to be present only at low frequencies. Pharmacological mechanism related to serious ADRs can occur in sensitive individuals, those with unusual kinetics, and in the presence of kinetic or occasionally dynamic drug interactions. In principle, such ADRs might be predictable from safety pharmacology testing, although it should be acknowledged that safety pharmacology testing is usually conducted in young adult healthy animals, conditions that may be suboptimal to detect such effects. Occasional nonpharmacological serious ADRs occur; these can be induced by direct chemical toxicity, hypersensitivity, or immunological mechanisms. Serious ADRs always limit the use of a new medicine by requiring warnings, precautions, and contra-indications; they can even preclude regulatory approval. Apart from preventing the development of NCEs likely to induce serious ADRs in larger clinical studies, a key contribution that can be made by non-clinical safety pharmacology is in the elucidation of the mechanisms responsible for these ADRs. Once the mechanism responsible for the ADR is known, it becomes possible to prepare soundly argued precautions and contra-indications.

When medicines are on the market, the actual incidence of serious ADRs is difficult to judge but clearly they occur with sufficient frequency to be a serious concern. An authoritative review[7] concludes that between 1 in 30 and 1 in 60 physician consultations result from ADRs (representing 1 in 30 to 40 patients). The same review concludes that 4 to 6% of hospital admissions could be the result of ADRs. Although there is debate over the number of deaths caused by ADRs — the figure of around 106,000 deaths per year in the USA is often quoted,[2,8] this has been suggested to be a gross overestimate and, for example, the U.S. FDA MedWatch system recorded 6894 deaths in 2000 (http://www.fda.gov/medwatch/index.html).

The frequency of serious ADRs can be very low (e.g., 0.25 to 1.0 cases of rhabdomyolysis per 10,000 patients treated with a statin[9]); however, when millions of patients are under treatment, this can generate substantial morbidity. Furthermore, ADRs may be due to (1) clinical error (e.g., misprescribing contra-indicated drugs) or (2) patient self-medication error — especially in the era of mass media communication and information. Although the frequency of these events can be very low, it is still necessary to investigate the pharmacological mechanisms driving these events. For example, the elucidation of the connection between drug-induced Torsades de Pointes (TdP), QT interval prolongation and hERG

TABLE 13.4
Evolution of the Main Safety Reasons for Drug Withdrawal over the Last 40 Years

Worldwide Withdrawal (121 Medicines)[10]	Percentage	The U.S. Withdrawal (95 Medicines)[7]	Percentage
Hepatic toxicity	26	Cardiovascular toxicity incl. Arrhythmias	19 (12)
Hematological toxicity	10	Neuropsychiatric effects/ abuse liability/dependency	12
Cardiovascular toxicity	9	Hepatic toxicity	9
Dermatological effects	6	Bone marrow toxicity	7
Carcinogenicity issues	6	Allergic reactions	6

potassium channel blockade has been considered as a major advance in this area and led to the rapid development of non-clinical *in vitro* screening assays of medium- to high-throughput capabilities. To better understand the main causes for ADR-related drug withdrawals, medicines withdrawn from either the U.S. or worldwide market were reviewed.[7,10] The principal reasons, presented in Table 13.4, highlighted the fact that several of these toxicities fall into the remit of safety pharmacology such as cardiovascular, immune, and central nervous systems–associated ADRs. The prominence of arrhythmias in Stephens' review probably reflects the recent interest in TdP-type arrhythmias.

13.2 Origin and Evolution of Safety Pharmacology

13.2.1 Regulatory Requirements

Prior to 1990, regulatory guidance on non-clinical organ function testing was limited. The U.S. and European regulations provided only general references to the evaluation of drug effects on organ system functions.[11–13] Organ function assessments included within investigational new drug applications (INDs) and registrations (NDAs) were inconsistent and often viewed as unimportant.[14,15] However, in Japan, the Ministry of Heath and Welfare (MHW), now referred to as the Ministry of Health, Labor, and Welfare, had promulgated comprehensive guidance for organ function testing as early as in 1975 (see Table 13.5). These guidances described which organ systems would be evaluated as a first-tier evaluation (List A studies) and made specific recommendations regarding study designs (including description of models, criteria for dose selection, and which endpoints would be included in the investigation). These guidances also described a second tier of studies (List B) to be conducted based on the significant findings in List A investigations (Table 13.5).[16] Because the Japanese guidance were the most comprehensive of their time, they became the *de facto* foundation for organ function testing throughout the pharmaceutical industry.[2,17,18] The organ function studies included in Lists A and B were intertwined with studies whose aim was to catalog additional pharmacological functions and activities (i.e., secondary pharmacology) in addition to the primary pharmacological function/activity. Kinter et al.[12] distinguished two subgroups of objectives embedded in the Japanese studies as safety and pharmacological profiling. This concept was enlarged upon the International Conference on Harmonization (ICH) safety pharmacology. Expert Working Group (EWG) to define three categories of pharmacology studies: primary and secondary pharmacodynamic studies, and safety

pharmacology studies (see Section 13.3[18,19] and Table 13.5). During the same period, European, U.S., and Japanese regulatory agencies prepared positions on general pharmacology/safety pharmacology in the form of guidance and concept papers.[18–21] Any complacency surrounding safety pharmacology was shattered in 1996 with the appearance of the first draft of a "Points to Consider" document on QT prolongation by the European Medicines Agency's Committee for Proprietary Medicinal Products (CPMP) and issued as an official document the following year.[22] One of the more controversial aspects of this document was the recommendation to incorporate screening of all noncardiac drugs for effects on cardiac action potential *in vitro*. The opposition to this document from the pharmaceutical industry arose partly because this recommendation wrong-footed the industry. The positive impact of the CPMP document was that it resuscitated safety pharmacology as a rigorous scientific discipline. In 1998, the MHW and the Japanese Pharmaceutical Manufacturers' Association proposed to the ICH Steering Committee the adoption of an initiative on safety pharmacology. This proposal was accepted and given the designation of Topic S7.

The origin of the term safety pharmacology is obscure. It first appeared in the draft guidances of the ICH M3 and S6 (see Table 13.5).[23,24] ICH S6 stated that "…the aim of the

TABLE 13.5

International Guidances and Draft Documents Referring to Physiological Functions as Relevant to Safety Pharmacology Assessment

Document and Source	Comments	Reference
In 1975, the JMHW in notes on applications for approval to manufacture new drugs requested the evaluation of the effects of physiological functions	To assess the effects of the test substance on the central nervous system, peripheral nervous system, sensory organs, respiratory and cardiovascular systems, smooth muscles including uterus, peripheral organs, renal function, and adverse effects observed in clinical studies	Anon.[27]
Guidelines for General Pharmacology Studies — Japanese Guidelines for Non-clinical Studies of Drugs manual 1995	Defines a list of studies to be conducted on all NCE (List A) and a list of studies to be conducted on a case-by-case basis (list B). Guidance to be superseded by ICH S7A and S7B	Anon.[16]
CPMP. Points to Consider: The assessment of QT interval prolongation by noncardiovascular medicinal products	Recommends a set of non-clinical, including *in vitro* assays, and clinical studies to address the risk for a noncardiovascular drug to slow cardiac repolarization and prolong the QT/QTc interval	Anon.[23]
ICH M3: Timing of Non-clinical Safety Studies for the Conduct of Human Clinical Trials for Pharmaceuticals	Provides guidance on the timing of safety pharmacology studies in relation to clinical development. Establishes that safety pharmacology studies should be conducted prior to first administration to humans	Anon.[24]
ICH S6: Preclinical Safety Evaluation of Biotechnology-Derived Pharmaceuticals	Defines the objective of safety pharmacology studies to reveal functional effects on major physiological systems (e.g., cardiovascular, respiratory, renal, and central nervous systems)	Anon.[23]
ICH S7A: Note for Guidance on Safety Pharmacology Studies for Human Pharmaceuticals	Provides the general framework for *in vitro* and *in vivo* safety pharmacology studies, including studies addressing the risk for a drug to slow cardiac repolarization	Anon.[25]
Recommendations from the Scientific Advisory Panel Subgroups on Hepatotoxicity: Hepatotoxicity of Health Products	Recommends clearance studies as part of the assessment of liver function	Anon.[28]

(*Continued*)

TABLE 13.5 (*Continued*)

Document and Source	Comments	Reference
ICH S7B: Note for Guidance on the Non-clinical Evaluation for Delayed Ventricular Repolarization (QT Interval Prolongation) by Human Pharmaceuticals	Describes a non-clinical testing strategy for assessing the potential of a test substance to slow ventricular repolarization. Includes information concerning non-clinical assays and integrated risk assessment	Anon.[42]
ICH E14: Note for Guidance on the Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Nonantiarrhythmic Drugs	Provides recommendations concerning the design, conduct, analysis, and interpretation of clinical studies to assess the potential of a drug to slow cardiac repolarization	Anon.[53]
Draft Guideline on the Non-clinical Investigation of the Dependence Potential of Medicinal Products	Provides guidance on the need for testing of dependence potential in animals during the development of medicinal products and indicate what type of information is expected as part of a MAA	Anon.[60]
Draft Guidance on Exploratory IND Studies	Provides guidance on non-clinical and clinical approaches that should be considered when planning exploratory IND studies in humans	Anon.[45]

Note: GLP = Good Laboratory Practice; ICH = International Conference on Harmonization; IND = Investigational New Drug; JMHW = Japanese Ministry of Health and Welfare; MAA = Marketing Authorization Application; NCE = New Chemical Entity.

safety pharmacology studies should be to reveal functional effects on major physiological systems (e.g., cardiovascular, respiratory, renal, and central nervous systems)….” The ICH S7 EWG began its work in the first quarter of 1999, and a harmonized safety pharmacology guidance was finalized and adopted by the regional regulatory authorities over 2000 to 2001.[25] The ICH S7A guidance describes the objectives and principles of safety pharmacology, differentiates tiers of investigations, establishes the timing of these investigations in relationship to the clinical development program, and introduces the requirement for Good Laboratory Practice (GLP) where applicable.[25,26]

The ICH S7 EWG extensively debated how to evaluate the potential of new drugs to produce a rare but potentially life-threatening ventricular tachyarrhythmia (TdP) in susceptible individuals.[29–33] The incidence of TdP with drugs that are targeted at noncardiac indications can be very low, hence the imperative need to find non-clinical surrogates is to identify those drugs with the potential to elicit TdP.[34–41] The controversial issue has been the accuracy of the non-clinical models to identify problematic drugs and how the generated data may be assimilated into an assessment of human risk.[17] Recognizing that the resolution would not be easily forthcoming, the ICH S7 EWG proposed to the ICH Steering Committee a new initiative to generate guidance for the assessment of the effects of drugs on cardiac ventricular repolarization. This proposal was accepted in November 2000 and was designated ICH S7B (see Table 13.5).[42] The guidance on safety pharmacology was finalized at the same meeting and was redesignated S7A (see Table 13.5).[25] Surveys of the pharmaceutical industry, regulatory agencies and members of the audience of the 4th Safety Pharmacology Society meeting, conducted 3 years after the implementation of the ICH S7A, concluded that the guidance has been successfully implemented in which GLP-compliant safety pharmacology “core battery” studies are usually performed prior to first administration to humans.[43] The approach is science-driven and specifies the use of robust and sophisticated *in vitro* and *in vivo* assays.

There are, however, some areas that require further refinement/clarification such as the specifics of study design including the selection of dose–concentration, choice of

species, modeling of the temporal pharmacodynamic changes in relation to pharmacokinetic profile of parent drug and major metabolites, use of an appropriate sample size, statistical power analysis, testing of human-specific metabolites, and demonstrating not only the model's sensitivity but also its specificity for predicting AEs in humans.[43] A year after the adoption of the ICH S7B, the U.S. Food and Drug Administration (FDA) and the Pharmaceutical Research and Manufacturers of America proposed to the ICH Steering Committee the adoption of a parallel initiative to prepare guidance on clinical testing of NCEs for their potential to prolong ventricular repolarization. This proposal was accepted as ICH E14 (Table 13.5). Following a recommendation from the ICH Steering Committee, the activities of the ICH S7B and E14 EWGs were aligned from 2003 onward. In May 2005, both guidances were finalized and due for implementation in November of the same year (see Table 13.5).[42,53] In a field where the scientific understanding and technological advances evolve rapidly, it would be important to insure that formal processes are in place to review and update guidance documents as need be.

13.2.2 Factors Influencing the Approach to Safety Pharmacology

In addition to regulatory requirements, several factors will influence the safety pharmacology strategy in any given pharmaceutical organization. The main factors are presented in Table 13.6. Thus, the pharmaceutical industry is beset with a number of significant challenges to achieve high-quality, high-throughput, and predictive safety pharmacology studies, during the early stages of the discovery process. Along with satisfying project demands, scientific safety questions, international regulatory guidances, and increased patient awareness, safety pharmacology is increasingly being used to unable informed decision making. One of the key factors influencing the approach to safety pharmacology is the rapidly evolving scientific and technological knowledge. Recent examples include the evolution (1) from convention labor intensive patch-clamp electrophysiology to medium throughput electrophysiology based platforms; (2) from manual measurements of the electrocardiogram (ECG) and blood pressure in conscious animals to semiautomated devices and analytical software; (3) from the manual counting of respiration rate to the direct quantification using whole-body plethysmography chambers; and (4) from the assessment of gastrointestinal (GI) function (gastric emptying, intestinal transit) using charcoal meals to the utilization of scintigraphy techniques.

TABLE 13.6

Non-Exhaustive List of Factors Influencing the Approach to Safety Pharmacology

Increase number, complexity, and stringency of regulatory requirements
Increase number and novelty of new chemical entities
Increase number and novelty of molecular targets and approaches (e.g., monoclonal antibodies, gene therapy)
Increase throughput of *in vitro* vs. *in vivo* assays
Increase risk identification initiatives during early discovery stages (e.g., "front loading" initiatives)
Increase awareness and application of the "3Rs" rule of animal usage and welfare
Increase patient awareness and expectations
Reduce supply of compound during the early discovery stages
Reduce availability of scientific and technical expertise in key areas (e.g., integrative physiology and pharmacology)
Reduce late-stage attrition
Reduce discovery and development time
Predictive value of *in vivo* and *in vitro* non-clinical assays with respect to human safety

13.3 Definition and Objectives of Safety Pharmacology Studies

During the course of the discovery and development of a drug, three types of pharmacology studies are to be conducted, namely primary, secondary, and safety pharmacology. Primary pharmacology studies are defined as those that "investigate the mode of action and effects of a substance in relation to its desired therapeutic target." On the other hand, secondary pharmacodynamic studies are defined as those that "investigate the mode of action and effects of a substance not related to its desired therapeutic target," whereas safety pharmacology studies "investigate the potential undesirable pharmacodynamic effects of a substance on physiological functions in relation to exposure in the therapeutic range and above."[20,25] Therefore, safety pharmacology studies are designed to investigate functional effects as opposed to morphological changes of an NCE.[46] Although pharmacology studies have been divided into subcategories, it is recognized that undesired functional effects (the domain of safety pharmacology) may be mediated via the primary or secondary pharmacological targets (see Table 13.3). Therefore, it is important to take a holistic approach while assessing adverse functional effects. The objectives of safety pharmacology studies are threefold: first "to identify undesirable pharmacodynamic properties of a substance that may have relevance to its human safety;" second "to evaluate adverse pharmacodynamic and pathophysiological effects of a substance observed in toxicology and clinical studies;" and third "to investigate the mechanism of the adverse pharmacodynamic effects observed and suspected."[25] Thus, the ICH S7A guidance objectives are primarily concerned with protecting clinical trials participants (volunteers and patients) and also patients that are receiving marketed products from any potential adverse effects of NCEs. This view is further supported by the clinician's perspective who sees the objectives of safety pharmacology as enabling (1) adapting the design (including parameters) of clinical studies; (2) preventing serious ADRs in early clinical trials; (3) providing guidance for setting up the doses/exposures in ascending dose tolerance clinical studies; (4) predicting the likelihood of unwelcome pharmacologically mediated ADRs that need monitoring during early clinical studies; and (5) reducing or eliminating serious ADRs in large-scale clinical trials and in clinical practice. Moreover, an unwritten objective may be to support business decisions mainly based on predictions of likely human safety profile or by identifying the risk in the early phases of drug discovery in order to design out, wherever possible, unwanted pharmacological activities.

13.4 Current Practices

13.4.1 ICH S7A/B

13.4.1.1 Core Battery, Follow-Up, and Supplemental Studies: Definitions and Expectations

Safety pharmacology studies have been subdivided into "core battery," "follow-up," and "supplemental" studies.[25] The "core battery" studies are aimed to investigate the effects of NCEs on the cardiovascular, respiratory, and central nervous systems that are considered as vital organ systems based on the fact that acute failure of these systems would pose an immediate hazard to human life. In some instances, based on scientific rational, the "core battery" may or may not be supplemented.[25] Additionally, ADRs may be either (1) suspected based on the pharmacological class, or the chemical class, or (2) identified

based on outcome from other non-clinical or clinical studies, pharmacovigilance, or from literature reports. When such potential ADRs raise concern for human safety, these should be explored in "follow-up" or "supplemental" studies. "Follow-up" studies are meant to provide a greater depth of understanding or additional knowledge to that provided by the "core battery" on vital functions. Moreover, "supplemental" studies are meant to evaluate potential adverse pharmacodynamic effects on organ system functions that have not been addressed in either the "core battery" or repeated dose toxicity studies when there is a cause for concern. The organ systems falling into this category may include, but are not limited to, the gastrointestinal (GI), renal, urinary, immune, endocrine, or autonomic nervous systems.

13.4.1.2 General Considerations and Principles

Since the pharmacological effects of an NCE depend on its intrinsic properties, the studies should be selected and designed accordingly. General considerations in the selection and design of these studies that assess potential ADRs are to include effects associated: (1) to the therapeutic class, (2) with members of a chemical class, (3) to nontarget mediated activities, and (4) to effects observed in previous non-clinical/clinical studies that warrant further investigation.

13.4.1.2.1 Species, Gender Selection, and Animal Status

The selection of an appropriate species is crucial in having an understanding of the molecular and biochemical comparison of the underlying controls on the physiological system(s) in the test species with those that are operative in humans. As an example, the dog and monkey are considered appropriate species for evaluation of drug effects on cardiac ventricular repolarization because of their dependence on potassium rectifying currents, IKr and IKs, to repolarize the ventricular myocardium; a similar dependence is known to exist in the human myocardium. On the other hand, the primary cardiac repolarizing currents in mice and rats rely on the outward potassium current, Ito. Thus, these species would not be appropriate for assessing the potential human risk posed by drugs that can affect cardiac IKr and IKs repolarizing currents. As another example, the rat does not possess a gall bladder; this raises important questions about the relevance of this specy for assessing GI function. On the basis of GI functional homology for humans, especially motility, gastric emptying, and pH value, particularly in the fasted state which is analogous to the conditions prevailing in many Phase I trials, the dog is perhaps a more relevant species.[47] Moreover, the dog appeared as a better predictor of clinical GI ADRs than the monkey for 25 anticancer drugs.[48] Although physiological similarity is an important requirement, it is only one of many factors that must be considered during species selection. In the absence of specific scientific reasons, the species in safety pharmacology studies are selected to maintain their consistency across study types (i.e., toxicology studies where information on metabolism and toxicokinetic are available).

Ideally, the gender selected for a study should include both male and females. However, it could be argued that non-clinical safety pharmacology studies should focused primarily on the gender that will be included in Phase I (i.e., usually male) or on the most sensitive gender for a given assay (e.g., female gender to assess the proarrhythmogenic potential).[49] It is important to note that these animals are most often healthy, which may be an important distinction from the patient population for which the drug is being developed. In many cases, the patient population may have an enhanced risk of demonstrating ADRs. To compensate for this possible limitation, the exposure of animals to the NCE and any major metabolites should explore large multiples (e.g., 100-fold) of the anticipated therapeutic concentration wherever feasible.

13.4.1.2.2 Dose or Concentration Selection, Route of Administration, and Duration of Studies

Doses selected for safety pharmacology studies are typically based on the criteria established in the ICH S7A guidance.[25] Doses should exceed those projected for clinical efficacy and at the upper limit be bound by (1) adverse pharmacodynamic effects in the safety pharmacology study; (2) moderately adverse effects in other non-clinical studies that follow a similar route and duration of dosing; or (3) limit of solubility/toxicity. In the absence of adverse effects, the maximum administrable dose can be used. If nonreusable animals enter the study, then the maximum tolerated dose may be appropriate. Most importantly, the doses/concentrations should establish the dose/concentration–response relationship of the adverse effect.

The route of administration of an NCE is typically the intended clinical route of administration. However, an alternative route may be used if this leads to an increase in systemic exposure of parent drug or major metabolites or if this alternative route satisfies another important objective of the study. For example, it is common to increase the exposure following inhalation administration by associating a subcutaneous administration of the NCE.

Data should be collected for a sufficient period of time to identify both the onset and recovery of effects should they be seen. In absence of knowing whether effects of the NCE will be observed in the non-clinical study, the timing of measurements will be based on available pharmacokinetic or toxicokinetic data collected in the same species, using the formulation and route that have been selected for the safety pharmacology investigation. On the basis of pharmacokinetic/toxicokinetic profile of the NCE, measurements will be made for a period that encompasses the maximal blood/plasma concentration (i.e., C_{max}) of parent drug and major metabolites, with recovery encompassing a period of at least five half-lives beyond the C_{max}. Moreover, a recent guidance document suggest that for all pivotal non-clinical safety studies that include a toxicokinetic evaluation, control samples should be collected and analyzed irrespective of the route of administration to confirm the validity of the study.[44]

For *in vivo* studies, the Latin-square crossover design can be adopted. A relatively small pool of animals can be reused to study vehicle and two to three dose levels of drug as long as an appropriate washout period (at least five half-lives) is permitted between each dose day. Most often, the studies follow single-dose administration, with an adequate washout period between doses when animals are reused. However, an NCE that exhibits a prolonged half-life (e.g., antibody vs. small molecule) may not allow for execution of this type of a Latin-square crossover study design within a practical timeframe. In this case, one may use a single-dose study design, allowing up to 1 or 2 weeks between doses and accepting that some accumulation of drug will occur with each subsequent dose, or a study of discrete groups of animals. In both cases, the use of a parallel negative control group (i.e., vehicle or placebo treatment) will be important to assess any temporal changes associated with the environment, human intervention, accommodation of animals, or the handling procedures.

The use of a negative control group (i.e., placebo or vehicle) is a requirement. In particular, effects of novel vehicles should be assessed since it is recognized that pharmacological properties of some vehicles may interfere with the activity of the NCEs and compromise with the interpretation of the results. The use of a positive control in each study may not be practical based on the need to demonstrate multidirectional changes in a multitude of parameters as possible outcomes of administration of the NCE. This concern has been addressed in the ICH S7A guidelines that does not require the use of a positive control in each *in vivo* study, but rather by demonstrating that the *in vivo* model has been fully characterized.[25] One situation, however, where a positive control may be

warranted is to rule out a suspected activity of the NCE by demonstrating the sensitivity and specificity of the experimental model. Additionally, the use of positive control is recommended for *in vitro* studies, where results of the NCE might be normalized to the effects of the positive control.[42]

13.4.1.2.3 Application of Good Laboratory Practice

Similar to other non-clinical safety studies that are required for registration of human pharmaceuticals, the GLP standards[26] apply to safety pharmacology studies,[25,42] with some exceptions as noted below. GLP adherence is expected for the safety pharmacology "core battery" studies, but the guidance acknowledges that aspects of the "follow-up" and "supplemental" studies, if they use unique methodologies, may not necessarily comply with GLP standards.[25,42] In these cases, however, the guidance recommends that the reconstruction of the study must be insured through documentation and archiving, and adequate justification is required for those aspects of the study not complying with GLP standards. Additionally, the impact of the noncompliance on the generated results and their interpretation should be acknowledged.[25,42,43]

13.4.1.2.4 Application of Statistical Analysis

The application of statistical methodology should be appropriate to optimize the design of the study to detect changes that are biologically significant for human safety,[25,50] while avoiding the unnecessary use of animals. The use of statistical power analysis allows the determination of optimal sample size for the detection of biologically relevant changes. For example, group sizes of 4 to 8 dogs are sufficient to detect, with an 80% chance, a 10 to 15% change in cardiovascular parameters (e.g., blood pressure, left ventricular pressure, QT interval, QTc interval).[51,187]

13.4.1.2.5 Testing of Isomers, Metabolites, and Finished Products

Another challenge is how and when to consider isomers, metabolites, and the actual finished product. In general, any parent compound and its major metabolite(s) that achieve, or are suspected to achieve, systemic exposure in humans should be evaluated. Assessment of the effects of major (i.e., >25% of the parent) human-specific metabolite(s), if absent or present only at relatively low concentrations in animals, should be considered.[25,52] This is of particular importance if the metabolite(s) is known to substantially contribute to the pharmacological actions of the NCE. *In vitro* or *in vivo* testing of the individual isomers should also be considered. Moreover, studies on the finished product are only necessary if the pharmacokinetics/pharmacodynamics is substantially altered in comparison to the active NCE tested previously.

13.4.1.2.6 Conditions under Which Safety Pharmacology Studies Are Not Necessary

The ICH S7A guidance makes provision for conditions under which safety pharmacology studies may not be necessary.[25] The conditions include: (1) locally applied agents where the systemic exposure or distribution to other organs or tissues is low, (2) cytotoxic agents of known mechanisms of action for treatment of endstage cancer patients, (3) new salts having similar pharmacokinetics and pharmacodynamics properties to the original NCE, and (4) biotechnology-derived products that achieve highly specific receptor targeting. In the latter example, the evaluation of safety pharmacology endpoints may be considered as part of the toxicology studies. However if the biotechnology-derived products represent a novel therapeutic class or are not achieving high selectivity, a more extensive evaluation of safety pharmacology studies should be considered.

13.4.1.2.7 Timing of Safety Pharmacology Studies in Relation to Clinical Development

The safety pharmacology "core battery" studies should be available prior to first administration in humans. Furthermore, "follow-up" and "supplemental" studies should also be available prior to first administration in humans if there are specific cause for concerns. During clinical development, additional studies may be warranted to clarify observed or suspected adverse effects in animals or humans. Finally, prior to product approval, effects on organ systems that are defined as part of "follow-up" and "supplemental" studies should be assessed, unless not warranted. Available information from toxicology or clinical studies can support this assessment and replace the need for stand-alone safety pharmacology studies.

13.4.2 Assessment of Vital Organ Functions

The following sections provide examples of approaches to assess drug effects on the cardiovascular, respiratory, and central nervous systems in compliance with the current and emerging regulatory guidance documents.[25,42,45,60]

13.4.2.1 Cardiovascular System

A comprehensive evaluation of the cardiovascular system includes an assessment of heart and vascular function and an evaluation of alterations in blood components. The heart functions as a "pumping unit" and may be impacted by effects of NCEs on the contractile elements of the myocardium or a loss in synchrony of the depolarizing wavefront that transits the myocardium (the conduction of an electrical impulse from the atrium to the ventricle) or repolarization of the myocardium (recovery of the myocardial cell to allow the propagation of the next depolarizing impulse). Effects of NCEs on vasculature may be manifested by a redistribution of flow to specific vascular beds generated by changes in vascular resistance. Effects of an NCE on the heart and vasculature may be mediated through direct actions on receptors, ion channels, transporters, enzymes or intracellular second-messenger systems or indirect effects on neurons, hormones, or normal physiological reflex mechanisms.

Measurements of heart function may include an evaluation of cardiac rate, output, and contractility. These indices of heart function are reflected in the heart rate (derived from either the pressure pulse or ECG), systemic blood flows, myocardial shortening, intra-cardiac pressures, cardiac wall thickness, and cardiac chamber size; the latter four parameters were measured during both systole and diastole. Electrical conduction from the atrium to the ventricle during depolarization and recovery of the myocardium during repolarization are evident in the ECG recorded from electrodes placed either on the body surface, just below the surface (subcutaneous), or in close proximity to the heart within the thoracic chamber. The durations and amplitudes of PR and QRS intervals of the ECG represent conduction through the atrium and ventricle, respectively, and QT interval encompasses both phases of ventricular depolarization and repolarization. In addition to measurements of the duration of these specific intervals, an investigator will also interrogate the ECG to identify any changes in the morphology that may be indicative of pharmacodynamic effects on discrete areas of the heart or drug-related pathology that may have resulted from acute, subchronic, or chronic exposure.

The integrated function of the vasculature and heart, as a closed circulatory system, supplies nutrients and oxygen to critical organs and removes metabolic wastes and carbon dioxide. This integrated system results from the careful control of cardiac output, arterial blood pressure (systolic and diastolic pressures; integrated to derive mean arterial pressure), and systemic vascular resistance, thereby maintaining blood perfusion through

organs that are critical for sustaining life. Arterial blood pressure is equivalent to cardiac output (times) systemic vasculature resistance. Thus, an NCE's effect on either cardiac output or systemic vasculature resistance may increase or decrease arterial blood pressure. For example, dilation of the systemic vasculature (reflected as a decrease in systemic vasculature resistance) without a change in cardiac output will result in a decline in arterial blood pressure. In contrast, an increase in cardiac output without a change in systemic vascular resistance will result in an increase in arterial blood pressure. In practice, cardiac output and arterial blood pressure are measured parameters and systemic vascular resistance is derived.

Effects of NCEs on the cardiovascular system may also be secondary to alterations in the cardiac or vascular microstructure resulting from NCE-induced cellular toxicity. In the same light, changes in the components that constitute the blood following acute, subchronic, or chronic exposure to an NCE may also be manifested. These important pathologic endpoints are routinely assessed in multiple-dose toxicology studies in rodent and nonrodent species. Cardiovascular pharmacodynamic endpoints that are incorporated in such studies may serve two roles: (1) to identify any underlying NCE-induced changes in systemic hemodynamic or cardiac function that could have produced pathologic findings (i.e., an enlarged heart and myocardial necrosis that is the result of a sustained increase in cardiac afterload) or (2) to identify whether functional changes accompany or precede the pathologic lesion (i.e., alterations in the ECG that is associated with, but may precede the onset of a myocardial lesion). Depending on the nature of the lesion and the time course over which it develops, the results may provide a biomarker that can be monitored in the clinic to indicate the eventual emergence of the pathologic lesion with continued dosing.

Cardiovascular ADRs are one of the most prominent issues of the pharmaceutical industry; in the last decade, the single most common cause of the withdrawal or restriction of the use of drugs that have been already marketed has been the prolongation of the QT interval associated with polymorphic ventricular tachycardia, or TdP (see Table 13.4).[39,54–56] TdP is typically not seen in clinical trials prior to registration of the drug. For terfenadine, the recognition of this rare event required extensive use and detailed monitoring from 1985 until 1998 before the drug was finally recalled because of 125 suspected drug-related deaths in the United States alone. Recent epidemiological retrospective analysis suggests that the incidence of drug-induced TdP might have been underestimated.[57] TdP has been linked to delayed cardiac repolarization, as manifested by a prolongation of the QT interval on the ECG. As a consequence, NCE-induced QT prolongation is generally considered as a surrogate marker for drug-induced TdP. Almost all compounds that prolong the QT interval and produce TdP in humans do so via inhibition of the rapid form of the delayed rectifier potassium current, IKr. During the last few years, significant advances have been made in our ability to test for effects of NCEs on the IKr current and a range of other cardiac ion channels. These *in vitro* assays, in conjunction with the *in vivo* cardiovascular/cardiac models (see Table 13.7), are able to detect the vast majority of compounds capable of prolonging the QT interval in humans. Thus, the risk of NCE-induced changes in cardiac repolarization in humans has been greatly reduced but not eliminated.[54] Other notable recent examples of cardiovascular ADRs include the increased incidence of heart valve regurgitation of the anorectic fenfluramine, in association with phentermine[58] and the withdrawal from the market place of the cyclooxygenase Type 2 inhibitor (Vioxx) for an unacceptable increased risk of myocardial infarction and stroke.[59]

Regulatory guidance for non-clinical cardiovascular safety pharmacology testing is given in the ICH S7A and B.[25,42] The effects of an NCE on blood pressure, heart rate, and the ECG should be evaluated. Furthermore, *in vivo*, *in vitro*, and *ex vivo* evaluations, including methods for (assessing) repolarization and conductance abnormalities, should

TABLE 13.7

Examples of Commonly Used Cardiovascular "Follow-Up" Studies

Physiological EndPoint	Methodology/Test	Reference
In vitro		
Disaggregated cells		
Repolarizing currents (e.g., IKs, IK1, Ito), depolarizing currents (e.g., INa) currents, ICa (whole cell patch-clamp)	Disaggregated cells ventricular myocytes; mouse atrial tumor cells (AT-1); immortalized cardiac muscle cells (HL-1)	Jost et al.;[65] Liu and Antzelevitch;[66] Jurkiewicz and Sanguinetti;[67] Li et al.;[68] Yang and Roden;[69] Banyasz et al.;[70] Xia et al.[71]
APD (whole cell patch-clamp)	Disaggregated cells	Davie et al.[72]
Myocyte contraction, Ca transients with fluorescent dyes	Disaggregated cells	Suetake et al.;[73] Cordeiro et al.;[74] Graham et al.;[75] Hamilton et al.[76]
Heterologous expression systems		
hERG current (whole cell patch clamp)	Mammalian cell expression (CHO, HEK-293, mouse L-cells, COS-7) and *Xenopus* oocytes expression system	Witchel et al.;[77] Zou et al.;[78] Cavero et al.;[79] Martin et al.;[80] McDonald et al.[81]
KvLQT1-minK current (whole cell patch-clamp)	Mammalian cell expression (CHO, HEK293)	Sanguinetti et al. [82]
Affinities/functional activities at relevant molecular targets	Pharmacological profiling — binding or functional assays at receptors, enzymes, ion channels, transporters	Wakefield et al.;[83] Bowes et al.[20]
Isolated tissues		
Vasoconstrictor/vasodilator assessment	Isolated tissues (e.g., arterial/venous rings)	Krasner et al.;[84] Lefer et al.[85]
Chronotropism and inotropism	Isolated atrium/papillary muscle/isolated hearts	Sugiyama et al.;[86] Voss et al.[87]
APD (microelectrode)	Purkinje fibres; papillary muscle; mid-myocardial (M-cell) wedge preparation	Abi-Gerges et al.;[88] Gintant et al.;[89] Lu et al.;[90] Antzelevitch et al.[91]
MAP, ECG parameters, proarrhythmia	Isolated intact hearts (Langendorff)	Hondeghem et al.;[92] Eckardt et al.[93,94]
Flows, pressures	Isolated intact hearts (working)	Miyoshi et al.;[95] Hill et al.[96]
Aggregation	Platelets or whole blood cell assays	
In vivo		
Hemodynamic and cardiac parameters: pressure (arterial, venous, ventricular, e.g., ventricular contractility, cardiac output), HR, peripheral resistance, ECG parameters, body temperature, flow	Conscious (restrained or telemetry) and anesthetized	Takahara et al.;[97] Sato et al.;[98] Nekooeian and Tabrizchi[99]
Ejection fraction (% of blood expelled from the ventricle during contraction), % fractional shortening (contraction of left ventricle), stroke volume	Echocardiography	Ozkanlar et al.;[100] Tsusaki et al.;[101] Hanton et al.[102]

(*Continued*)

TABLE 13.7 (*Continued*)

Physiological EndPoint	Methodology/Test	Reference
(volume of blood ejected per contraction), d*p*/d*t* (an index of contractility), valve insufficiency or regurgitation		
Regional blood flow	Flow probes, Doppler, microspheres	Wakefield et al.[103]
Proarrhythmia	Conscious and anesthetized	Eckardt et al.;[93] van der Linde et al.;[104] Sugiyama et al.;[105] Chiba et al.[106]
Denervated autonomic nervous system	Denervated autonomic nervous system; anesthetized pithed rat	Collister and Osborn;[107] AbdelRahman[108]
Autonomic nervous system (baro-, mechano-, and chemoreflex)	Heart rate variability, nerve recordings, bar-mechano reflex curves, tilt table	Verwaerde et al.;[109] Harada et al.;[110] Mangin et al.[111]
Orthostatic maneuvers	Tilt table	Bedford and Dormer;[112] Humphrey and McCall[113]
Coronary reperfusion	Transient coronary artery ligature (total vs. partial occlusion)	Ytrehus[114]
Myocardial infarct size	Permanent coronary artery ligature	Wu and Lima[115]
Thrombosis	Models of arterial or venous thrombosis	Shebuski et al.[116]

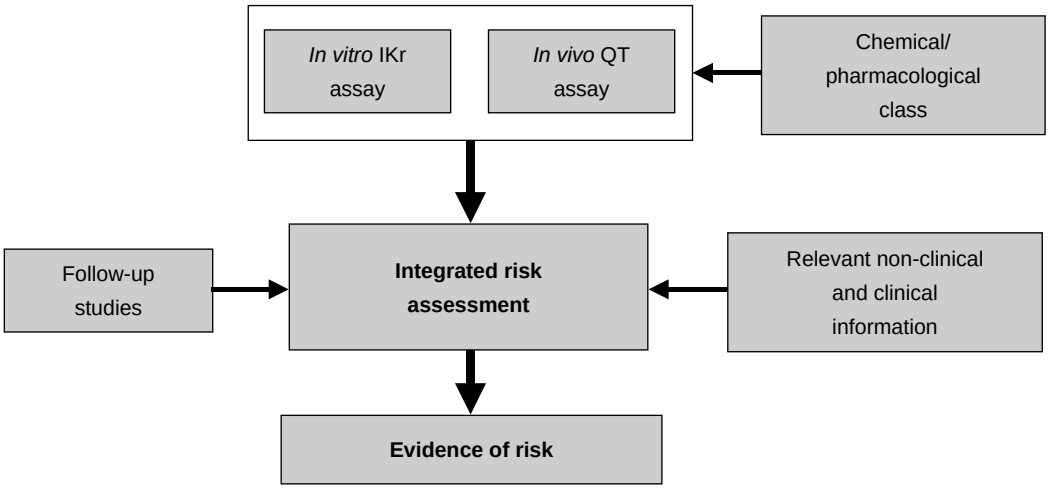


FIGURE 13.1
Component elements of the testing strategy for assessing risk for delayed ventricular repolarization and QT interval prolongation. (*Source*: Adapted from Anon., CPMP/ICH/423/02, 2005.)

also be considered. The evaluation of drug effects on the cardiovascular system can be conducted using a range of *in vivo* or *in vitro* techniques, and species (Table 13.7). The ICH S7B, however, specifically requires an *in vitro* IKr assay and an *in vivo* QT assay to be made available to the regulators (Figure 13.1).

13.4.2.1.1 In Vivo *QT Assay*

The aim of an *in vivo* QT assay is to measure indices of ventricular repolarization such as the QT interval. This assay can be designed to meet the objectives of both ICH S7A and S7B.[25,42] The development of telemetry techniques in conscious animals has had a major impact on the conduct of *in vivo* cardiovascular safety pharmacology studies. The telemetry technique permits a continuous collection of a range of physiological parameters, including heart rate, pressures (e.g., arterial, venous, pleural, left ventricular), ECG (including the QT interval), and body temperature over longer periods of time in undisturbed animals. Thus, drug effects can be studied under physiological conditions using the clinical route of administration. Alternatively, anesthetized animals can be used under conditions where (1) the compound is poorly tolerated in the chosen species (e.g., due to emesis or tremor), (2) its bioavailability/exposure is expected to be low, or (3) insufficient information is known about the NCE at the time of the evaluation. Although the effect of anesthesia on the NCE under study may be unknown (e.g., drug–drug interaction), it is generally accepted that effects seen in anesthetized animals are qualitatively comparable to those detected in conscious animals.

13.4.2.1.2 In Vitro *IKr Assay*

The aim of an *in vitro* IKr assay is to evaluate the effects of NCE on this ionic current. This evaluation can be performed in cardiac myocytes or cell lines expressing hERG, the α-subunit of the IKr channel protein. Screening of NCEs for hERG inhibitory activity early in the drug discovery process may help to make business decisions on hERG data alone that ultimately may reduce attrition of NCEs in later non-clinical or clinical development stages. Screening during the early phases of the discovery process implies the ability to test a large number of compounds in a short period of time with minimal effort in terms of resources. Several approaches are currently being used. In binding assays, the displacement of a radioactively labeled channel antagonist (e.g., [^{3}H]dofetilide) by the compound under investigation is measured.[61] Rubidium (Rb) flux assays rely on the high permeability of Rb^+through voltage sensitive K^+ channels.[62] Fluorescence assays make use of voltage-sensitive dyes, which measure the membrane potential of a living cell.[63] All these tests suffer from measuring the effects on the hERG channel indirectly and are therefore prone to artifacts. For the last two decades, the patch-clamp technique has been regarded as the gold standard, although the manual handling of the patch-clamp systems is labor-intensive. During the past few years, the introduction of new medium-/high-throughput patch-clamping technologies has significantly increased the ability to detect and design hERG-like properties.[64]

In addition to the "core battery" assays described above, the ICH S7A guidance recommends "follow-up" studies to investigate the effects on cardiac output, ventricular contractility, vascular resistance as well as the effects of endogenous and exogenous substances on the cardiovascular responses. A nonexhaustive list of *in vitro*, *ex vivo*, and *in vivo* cardiovascular models is presented in Table 13.7 that can be used to investigate known or suspected issues. With respect to assessing the potential for an NCE to slow ventricular repolarization and prolong the QT interval, "follow-up" studies can be used to understand the basis of discrepancies among non-clinical studies and between non-clinical and clinical studies (Figure 13.1).

13.4.2.2 Respiratory System

The respiratory system can be divided functionally into a pumping apparatus and a gas exchange unit (see Figure 13.2).[117] The pumping apparatus includes those components of the nervous and muscular systems that are responsible for generating and regulating breathing patterns, whereas the gas exchange unit consists of the lung with its associated

airways, alveoli, and interstitial area that contains blood and lymph vessels and an elastic fibrous network. The ICH S7A guidance recommends that respiratory rate and other measures of respiratory function (e.g., tidal volume or hemoglobin oxygen saturation) should be quantified using appropriate methodologies as part of the "core battery" studies. The assessment of lung mechanics (e.g., airway resistance, lung compliance) has been relegated to the rank of "follow-up" investigations and therefore is not required prior to first administration to humans (unless there is a cause for concern).[25]

In some respects, part of the content of the ICH S7A guidance regarding respiratory function monitoring and respiratory testing is inaccurate and misleading.[25,43,130] Therefore, to effectively protect clinical trials participant from respiratory ADRs while avoiding unnecessary animal usage, it is recommended that the effects of NCEs on respiratory function should include both ventilatory and mechanical function assessments.[43,130] The primary test for evaluating the pumping apparatus is the measurement of ventilatory changes in conscious animals. To characterize ventilatory patterns, the measurement should include at least the respiratory rate, tidal volume, and minute volume. Additionally, to investigate for potential mechanisms, the measurement should also include the inspiratory flow, expiratory flow, and fractional inspiratory flow. "Follow-up" evaluations should include tests for detecting the occurrence of hypo- or hyperventilation syndromes and for distinguishing central from peripheral effects. Functional changes in the gas unit, or lung, are evaluated by measuring changes in the mechanical properties of the lung. The primary test for evaluating the function of the lung should include measurements of the lung resistance and compliance. Dynamic measurements of the resistance to lung airflow and lung compliance are the preferred tests, as they can be used to simultaneously measure both the ventilation and lung

TABLE 13.8

Examples of Commonly Used Respiratory "Follow-Up" Studies

Physiological EndPoint	Methodology/Test	Reference
Ventilatory disorders		
Hypo-/hyperventilation	$PaCO_2$ *in vivo*, respiration rate, tidal volume, minute volume	Murphy et al.[117]
Central or peripheral mechanism	Responses to CO_2 and NaCN *in vivo*	Murphy et al.[118]
	Neuronal activity *in vitro*/ *in vivo*	Jackson et al.;[119] Widdicombe[120]
Obstructive disorders		
Airway narrowing	Forced manoeuvres (FVC, FEV, IC) *in vivo*; airway resistance *in vivo*; airway resistance *in vitro* (tracheal rings)	Diamond et al.;[121] Mauderly;[122, 124] O'Neil et al.;[123] Douglas et al.;[125] Kenakin;[126] Murphy et al.[142]
Restrictive disorders		
Interstitial thickening	Lung compliance *in vivo*	Diamond et al.;[121] Mauderly;[122, 124]
Hemorrhage	Ciliary beat assays *in vitro*	Chilvers et al.[127]
Cellular infiltration	Mucus level assays *in vitro*	Khawaja et al.[128]
Abnormal surfactant production	Alveolar Type II cell surfactant secretion assay *in vitro*	Dietl et al.[129]
Affinities/functional activities at relevant molecular targets	Pharmacological profiling — binding or functional assays at receptors, enzymes, ion channels, transporters	Bowes et al.;[20] Wakefield et al.[83]

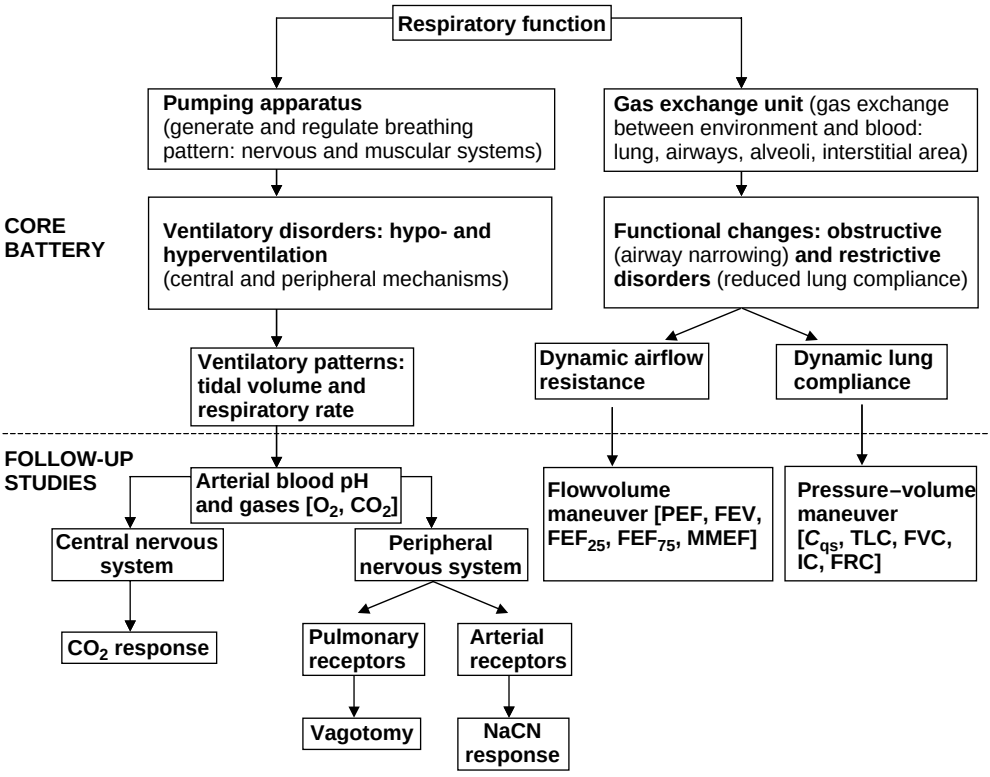


FIGURE 13.2
Recommended approach to the evaluation of drug-induced effects on respiratory function. "Core battery" tests are conducted to detect the occurrence of a functional change on both ventilatory parameters and lung mechanics, while "follow-up" studies are conducted to investigate mechanisms, further characterize effects, or define liability associated with a functional change. PEF = peak expiratory flow; FEV = forced expiratory volume; FEF25 = forced expiratory volume at 25% of FVC; FEF75 = forced expiratory volume at 75% of FVC; MMEF = maximum midexpiratory flow; Cqs = quasi-static respiratory compliance; TLC = total lung capacity; FVC = forced vital capacity; IC = inspiratory capacity; FRC = functional residual capacity. (*Source*: Adapted from Murphy, D.J., *Fund. Clin. Pharmacol.*, 16, 183–196, 2002.)

function parameters repeatedly in conscious animals.[130] "Follow-up" evaluations should be performed to determine the site and extent of an obstructive disorder, to confirm the presence of a restrictive disorder, and to characterize effects on forced expiratory airflows and lung capacities (see Table 13.8 and Figure 13.2). Although the ICH S7A guidance refers primarily to *in vivo* studies, *in vitro* approaches can be deployed to address specific respiratory endpoints.[83]

13.4.2.3 Central Nervous System

Most of the adverse drug effects relating to the nervous system impact on the quality of life rather than the risk to life (e.g., lethargy, anorexia, insomnia, personality changes, and nausea). There are, however, some serious life-threatening adverse effects involving the nervous system (e.g., loss of consciousness and convulsions). Some of these reflect the fact that the nervous system controls the other two vital organ systems for that CNS impairment could be fatal (e.g., decreased respiratory drive leading to respiratory arrest; and decreased sympathetic outflow leading to cardiovascular collapse). The nervous system adjusts the function of the other acutely vital organ systems according to current and long-term

requirements of the organism. Therefore, drug effects on cardiovascular and respiratory functions can be mediated via a direct action within the central nervous system (CNS), or via sensory nerve endings located in the cardiovascular and pulmonary systems. Some CNS adverse effects can be indirectly life threatening. For example, drowsiness, cognitive impairment, motor coordination, dizziness, involuntary movement, and visual disturbances can all affect driving performance; moreover, depression and personality changes can lead to suicidal tendencies. As an illustration, the number of deaths in the United States between 1984 and 1996 in patients receiving terfenadine was 396, a proportion of which were attributed to sudden death resulting from TdP. This overall low incidence of fatalities is nevertheless a significant improvement over the first generation of antihistamine drugs that have been suspected to be responsible for significant fatalities in car accidents resulting from their sedative effects.[131,132]

The ICH S7A guidance states that effects of an NCE on "motor activity, behavioral changes, coordination, sensory/motor reflex responses, and body temperature should be evaluated. For example, a functional observational battery (FOB), modified Irwin's test, or other appropriate test can be used."[25,133–135] Whereas the Irwin's test was introduced in the pharmaceutical industry initially as a rapid psychotropic screening procedure for use in mice, the FOB arose from neurotoxicity testing in rats in the chemical and agrochemical industries and has been adapted and adopted for use in safety pharmacology testing of NCEs.[136–138] The FOB is a systematic evaluation of nervous system function in the rat, comprising more than 30 parameters and covering autonomic, neuromuscular, sensorimotor, and behavioral domains.[137]

"Follow-up" studies can include behavioral pharmacology, learning and memory, ligand-specific binding, neurochemistry, visual, auditory and electrophysiology examinations, etc. The ICH S7A guidance does not distinguish between CNS vs. non-CNS targeted drugs, although a higher incidence of CNS side effects would be expected for CNS-targeted drugs compared to non-CNS-targeted drugs; hence the number of "follow-up" studies is likely to be greater in the former case. In a recent survey, it was recommended that there would be a more extensive nonclinical assessment of CNS function.[43] The recommendations were based on the fact that some CNS adverse events could not necessarily be predicted from an Irwin's or FOB test (e.g., pro-/anticonvulsive potential, abuse potential, and headache).[43] Table 13.9 summarizes some *in vivo* and *in vitro* physiological functions that can be evaluated as "follow-up" studies on a cause-for-concern basis and the associated methodologies. Although the ICH S7A guidance refers primarily to *in vivo* studies, *in vitro* approaches can be deployed to address specific nervous system endpoints (for review see Wakefield et al.[83]). Of interest is the recent release for consultation of a draft guidance on the non-clinical investigations that might be required to assess the dependence potential of CNS-active NCEs.[60] The proposed generic approach is presented in Figure 13.3.

13.4.3 Assessment of Non-Vital Organ Functions

The ICH S7A guidance states that "supplemental" studies are meant to evaluate potential adverse pharmacodynamic effects on organ systems functions that are not acutely essential for the maintenance of human life and not addressed by the "core battery" or repeated dose toxicity studies when there is a cause for concern.[25] Examples of physiological functions that fall into that category include, but are not limited to, the renal/urinary, immune, GI, endocrine and autonomic nervous systems. This section focuses on the renal and GI systems based on their potential impact on the clinical development program.

TABLE 13.9
Examples of Commonly Used CNS "Follow-Up" Studies

Physiological Endpoint	Methodology/Test	Reference
In vivo		
Nociception	Tail flick, hot plate, plantar test	Eddy and Leimbach[139]
Convulsion	Electric or chemically induced convulsion	Krall[140]
Locomotor activity	Interruptions of photoelectric beams, activity wheels, changes in electromagnetic fields, Doppler effects video image analysis, telemetry	Reiter and McPhail[141]
Motor coordination	Rotarod; beam walking	Dunham and Miya;[143] Abou-Donia et al.;[144] Jolkkonen et al.[145]
Auditory function	Startle reflex; brainstem auditory evoked response	Redfern et al.;[2] Wakefield et al.;[83] Herr et al.[146]
Visual function	Ophtalmoscopy; optokinetic	Prusky et al.[147]
Anxiety	Elevated plus maze	Redfern and Williams[148]
Sleep induction	Loss of righting reflex; electroencephalogram	Porsolt et al.;[138] Jouvet[149]
Learning and memory	Passive avoidance task; Morris water maze; radial maze; operant behavior tasks	Bammer;[150] Glick and Zimmerberg;[151] Morris;[152] Olton;[153] Dunnett et al.[154]
Electrophysiological examinations	Electroencephalogram *in vivo*	Itil[155]
Drug dependence	Physical withdrawal phenomena; drug preference tests	Meert;[156] Goudie et al.[157]
Drug abuse	Place preference tests; drug discrimination tests; self administration procedures	Brady and Fischman;[158] Lal;[159] Schechter and Calcagnetti[160]
In vitro		
Electrophysiological examinations	Evoked potentials *in vitro* from brain slices	Fountain and Teyler[161]
Affinities/functional activities at relevant molecular targets	Pharmacological profiling — binding or functional assays at receptors, enzymes, ion channels, transporters	Bowes et al.;[20] Wakefield et al.;[83] Fliri et al.[162]

13.4.3.1 Gastrointestinal System

The GI tract, essentially a 10-m-long muscular tube extending from mouth to anus, is target to many clinical ADRs ranging from minor nonlife-threatening (e.g., tooth discoloration) to severe and life threatening (e.g., perforated ulcer).[163,164] Between the two ends of this severity spectrum range a multitude of common, yet poorly understood, drug-induced disturbances of GI function that negatively impact patient safety, compliance, quality of life, and clinical benefit. The impact of GI ADRs in terms of drug withdrawals from the market has been minimal, with only one example, pirprofen, in the period from 1960 to 1999;[10] however, drugs with labeling restrictions are commonplace, for example Lotronex.[165]

GI ADRs are some of the most frequently reported in all phases of clinical drug development and for marketed products, as illustrated by the 700 drugs that are implicated in

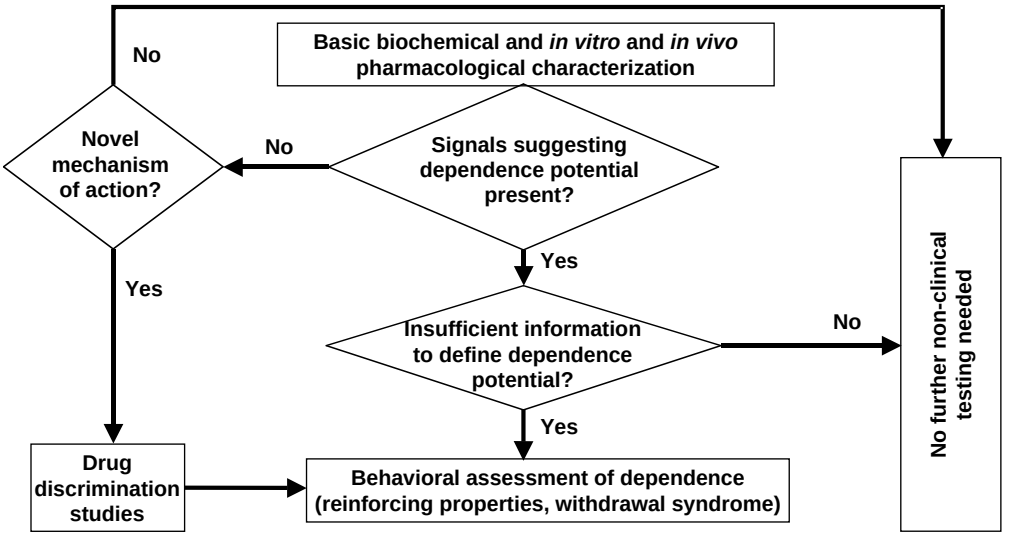


FIGURE 13.3
Proposed non-clinical testing strategy for assessing dependence potential of CNS-active NCEs. (*Source*: Adapted from Anon., EMEA/CHMP/SWP/94227/2004, 2005.)

causing diarrhoea.[166] Furthermore, GI ADRs account for approximately 18% of all reported clinical adverse drug reactions and 20 to 40% of those in hospitalized patients.[167] Given, however, that symptoms of disturbed GI function are encountered in everyday life, the actual incidence of drug-related effects is most likely extensively underreported. Most of the reported ADRs are functional in nature (nausea, vomiting, dyspepsia, abdominal cramps, and diarrhea or constipation) with a fewer number related to lesions (e.g., ulceration) or enhanced susceptibility to infection (e.g., pseudomembranous colitis).[168] Of these, it is estimated that approximately 80% are Type A-predictable pharmacological reactions.[169]

In addition to safety concerns, drug-related inhibition or enhancement of GI motor function can also lead to the alteration of a drug's pharmacokinetic profile owing to alterations in residency time of a drug in its site of absorption. Changes in plasma exposure of orally administered drugs, for example, aminophylline following application of an inhibitor of gastric motor function, for example, propantheline are well documented.[170]

The GI system is responsible at its most basic level for providing a continual supply of water, electrolytes, minerals, and nutrients. This is achieved by a myriad of specialized cells and coordinated interplay of motility, secretion, digestion, absorption, blood flow, and lymph flow. These components are under elaborate control of the central and enteric nervous systems, endocrine and paracrine regulation of hormones. The highly complex nature of GI function is clearly illustrated by the estimate that 80 to 100 million neurons exist within the enteric nervous system, a number comparable to that found within the spinal column, hence described as a "second brain."[171]

ICH S7A lists examples of GI parameters that can be measured; for example, gastric secretion, GI injury potential, bile secretion, transit time *in vivo*, ileal contraction *in vitro*, gastric pH, and pooling. The list of parameters is by no means comprehensive and s leans toward assessing function rather than histopathology; this inclination reflects the distribution of clinical GI ADRs, but functional assessment should not predominate at the expense of histopathology.

Harrison et al.[172] reviewed the methods and techniques capable of assessing specific changes in GI function at the membrane, cell, and whole animal levels. Membrane-based studies record the uptake of solutes and electrolyte transport, assessing the effects of NCEs

on transepithelial GI transport and flux. Such methods lend themselves to permeability, immunocytochemistry, morphology, and molecular biology techniques. Isolated cells from the GI tract or cultured cell lines provide knowledge of regulation and function at a cellular level, while motility pattern, taken from *in vivo* or from biopsies, provides information at a more integrated level. In anesthetized animals, ligated segments of the intestine can be infused with NCEs, providing information about absorptive and secretory processes important for the treatment of diarrhea. Computer simulations and modeling are used to simulate the dissolution, absorption, distribution, metabolism, and excretion properties of NCEs in the human GI tract, thereby replacing to some extent animal testing. Finally advances in the field of imaging, combined with endoscopy, have resulted in a wireless capsule, allowing the inspection of the GI tract anatomy and pathology without surgical intervention. One important element omitted from both ICHS7A and the review by Harrison et al.[172] of particular relevance to anticancer drugs is nausea and vomiting. Such omission may reflect the complexity of the symptom and that there are no non-clinical models able to relay nausea. There is, however, a growing understanding about the causes of nausea and vomiting[173,175] and non-clinical dog and ferret models amenable for safety pharmacology testing.[176,177] It is becoming increasingly evident that in addition to traditional *in vivo* techniques, safety pharmacologists could in the foreseeable future routinely incorporate membrane, isolated tissue, and endoscopy techniques for GI tract testing of NCEs.

13.4.3.2 Renal System

The kidney is a privileged target for toxic agents because of its physiological and pharmacokinetics properties. It receives the largest amount of blood per gram of tissue among any other organ and therefore it is more exposed to exogenous circulating NCE than many other organs. Moreover, tubular mechanisms of ion transport acts to facilitate drug entry into renal tubular cells. From a pharmacokinetic perspective, the kidney is involved in filtration, excretion, and reabsorption of NCEs. The kidney concentrates urine so that intratubular drug concentration may be much higher than plasma concentration, and finally the kidney has a high metabolic rate. Several drugs/drug classes are associated with nephrotoxicity (e.g., antibiotics, nonsteroidal antiinflammatory drugs, immunosuppressors, angiotensin-converting enzyme inhibitors, chemotherapeutic drugs, and fluorinated anesthetics).[178,179] Deterioration of renal function over a period of hours to days results in the failure to excrete nitrogenous waste products and the inability to maintain fluid and electrolyte balance. Acute renal failure owing to toxic or ischemic injury is a clinical syndrome referred to as acute tubular necrosis, and a common disease with high overall mortality (~50%).[11,180,181] There are a number of risk factors for acute renal failure that should be kept in mind during the development of an NCE, especially in clinical trials: (1) patient-related risk factors such as age, sex, race, preexisting renal insufficiency, specific diseases, sodium retaining states; dehydration and volume depletion, sepsis, shock; (2) drug-related risk factors such as dose, duration, frequency, form of administration, repeated exposure; and (3) drug–drug interaction such as the associated use of NCEs with added or synergistic nephrotoxic potential. The challenge while assessing renal function results from the ability of the kidney to adapt to increases in single nephron glomerular filtration rate that consequently tend to mask renal injury until a considerable amount of kidney parenchyma is irreversibly lost. Such impairments, although not detectable in normal populations, can be quite significant in susceptible individuals. As a consequence, the use of animals that have been made more sensitive to functional effects should be considered (e.g., salt depletion, dehydration, coadministration of pharmacological agents, and unilateral nephrectomized animals).[11]

The ICH S7A states that the effects of the NCE on renal parameters should be assessed; for example, urinary volume, specific gravity, osmolality, pH value, fluid/electrolyte balance, proteins, cytology, and blood chemistry determinations such as blood urea

nitrogen, creatinine, and plasma proteins can be used.[25] Most of these parameters are measured in rodent toxicology studies, although usually not after the first day of administration.

From a scientific point of view, glomerular filtration rate (inulin or creatinine clearance) is the best global estimate of renal function. Other parameters of interest include, but are not limited to:

1. Fractional excretion of electrolytes (i.e., sodium, potassium, and chloride)
2. Renal blood (or plasma) flow (*p*-aminohippuric acid clearance or ultrasonic transit time flowmetry)
3. Enzymuria (which could allow differential location of NCE-induced injuries, for example, alanine-aminopeptidase; γ-glutamyl-transferase; trehalase originating from the brush border; β-glucoronidase, *N*-acetyl-β-D-glucosaminidase, acid phosphatase, β-galactosidase originating from the lysosomes and lactate dehydrogenase, leucine aminopeptidase, β-glucosidase, fructose-1,6 biphosphatase, and pyruvate kinase originating from the cytosol)
4. Proteinuria (including albuminuria indicative of increases in the permeability of the glomerular capillary wall; β-2 microglobulin indicative of an impairment in tubular reabsorption)
5. Glucosuria (which may be indicative of proximal tubular damage if the serum glucose concentration is within the normal physiological range)
6. Diuresis
7. Concentrating ability of the kidney (measurement of urine osmolality assessed following withdrawal of food and water for 24 h; free water clearance)

Assessment of renal function is not required prior to first administration to man or even during clinical development; however, based on the potential implications of acute renal failure and the challenges in assessing it in normal healthy animals or humans, it would make sense to consider a proper assessment of renal function prior to first administration to humans.

13.4.4 Integration of Safety Pharmacology Endpoints in Toxicology Studies

Back in the late 1970s, Zbinden wrote that the adverse drug reactions, which the standard toxicological test procedures do not aspire to recognize, include most of the functional side effects. Clinical experience indicates, however, that these are much more frequent than the toxic reactions owing to morphological and biochemical lesions.[182] The commonly heard arguments against inclusion of functional tests in toxicology studies are the facts that (1) it is not a regulatory requirement; (2) the design, conduct, reporting, and interpretation of the tests require specialized scientific and technical expertise; (3) the increased handling may cause changes in stress hormones, which may influence the pathology; and (4) the tests are cost and labor-intensive for a limited value in return. However, several elements have argued in favor of the inclusion of functional endpoints in toxicology studies.[183] For example, some toxic responses are of a purely functional nature and are not accompanied by morphological lesions (e.g., arrhythmia and seizures); functional toxicity often occurs much earlier and at lower doses than those necessary to induce pathological organ damage; most pathological lesions are secondary to a functional disturbance (e.g., vasoconstriction, tachycardia, and endocrine responses).[182]

Table 13.10 illustrates examples of functional endpoints that have been successfully incorporated into toxicology studies. The examples cover primarily, but are not limited to, the vital organ systems (i.e., cardiovascular, respiratory, and central nervous systems). The repeat-dose toxicology studies offer an ideal opportunity to compare the pharmacological

TABLE 13.10

Examples of Measurement of Safety Pharmacology Endpoints in Toxicology Studies

Physiological Function	Parameters	Technique/ Method	Species	Reference
Respiratory	Respiration rate, inspiratory and expiratory times, tidal volume, minute volume, peak inspiratory, and expiratory flows	Whole body pletysmography	Rat	McMahon et al.[185]
Cardiovascular	ECGs intervals durations, amplitudes, and morphology	Surface electrodes using noninvasive telemetry in freely moving animals or conventional recordings in restrained animals	Dog	Hunter et al.;[186] Tattersall et al.;[187] Detweiller[188]
Cardiovascular	Heart weight, wall thickness, left ventricular (LV) and right ventricular (RV) end-diastolic and end-systolic lumen volumes (EDV and ESV, respectively), cardiac output (CO), heart rate, and LV diastolic filling pressure	Magnetic resonance imaging	Dog	Opie[189]
Cardiovascular	As above	Echocardiography	Dog, rat	Hamlin[190]
Nervous system	20–30 neurobehavioral endpoints	Functional observational battery	Rat (or mice) or dog	LeBel et al.;[191] Horner et al.[192]
Motor coordination	Time spend on the rotarod	Rotarod	Rat	Chapillon et al.[193]
Auditory	Amplitudes and durations of waveforms	Brain auditory evoked response	Dog	Holliday et al.[194]
Auditory	Startle reflex	Prepulse modulation	Rat	Colley et al.[195]
Visual	Opthalmoscopy	Ophthalmoscope	Dog	
Visual	Pupil diameter	Pupil diameter to dark/light stimuli	Rat	Redfern et al.[196]
Nociception	Time to withdraw tail	Tail flick	Rat	Horner et al.[192]
Muscle strength	Force	Grip strength	Rat	Horner et al.[192]
General activity	Distance traveled	Locomotor activity	Rat	Horner et al.[192]
Nervous system	Conduction time	Nerve conduction	Rat	Horner et al.[192]

(*Continued*)

TABLE 13.10 (*Continued*)

Physiological Function	Parameters	Technique/ Method	Species	Reference
Renal	Water intake, food intake, body weight, urine volume, urinary excretion of Na, K, Cl	Metabolism cage	Rat	
Gastrointestinal	Feces weight, aspect	Clinical observation or metabolism cage	Rat, dog	

effects of repeated dosing with single-dose responses. Some assessments are incorporated routinely as per regulatory requirements (e.g., assessment of renal function using metabolism cage; and assessment of cardiac electrical activity; Table 13.10). Some functional measures can be "bolted-on" without affecting the animals (e.g., home cage locomotor activity; Table 13.10) or with minimal impact (e.g., neurobehavioral assessment using a functional observational battery test; Table 13.10). Some functional endpoints can be incorporated on a case-by-case basis if a cause for concern arises from the primary, secondary or safety pharmacology studies, toxicology studies or clinical observations, for example, the measurement of brainstem auditory evoked responses (BAER) in conscious-restrained beagle dogs for a drug suspected to affect the auditory function (Figure 13.4); or the measurement of ventilatory parameters for a drug suspected or known to affect the respiratory function or lung morphology. Finally, under some circumstances where the functional measurements would interfere with the primary goal of the toxicology study, satellite groups of animals can be included (e.g., telemetered animals to measure blood pressure).

It is important to highlight that functional toxicology is not an alternative to acute dose safety pharmacology studies but it is complementary. The main reasons are as follows: (1) functional measurements during repeat-dose studies are assessing the "effects" rather than "responses" to drugs in animals that may have some degree of multiple organ impairment owing to repeated drug exposure; (2) the experimental conditions may be difficult to optimize, for the variable being measured, there may be a large number of animals to measure from in a short period of time, and other measurements are also being made; (3) a tolerance may develop to the drug response, for example, several classes of drugs/drug classes undergo a diminution of their initial response following repeated dosing (e.g., benzodiazepines, ethanol, nicotine, morphine, and β_2-adrenoceptors). This is an adaptive mechanism, occurring at molecular, cellular, and system levels related to the pharmacological target. It can also involve the induction of P450 enzymes. As summarized by Haefely, "Some form of adaptive syndrome is the inevitable consequence of the reciprocal interaction between most or all classes of drugs and the organism."[184] Functional endpoints can be integrated into regulatory toxicology. Technical progress in noninvasive methodology and refined measurements for pharmacological parameters and standardization of study design allow the incorporation into regulatory toxicology studies today and probably even more in the future. The limitation of conducting pharmacological measurements in regulatory toxicology studies should be acknowledged. Safety pharmacology studies should complement toxicity studies in terms of choice of species and dose regimen. Ethical consideration of animal usage, especially dogs and monkeys, can only be justified in the future when more clinically relevant data can be gained from fewer *in vivo* studies. Multidisciplinary cooperation between pharmacology, pharmacokinetics, and toxicology will lead to the refinement and reduction of *in vivo* studies when functional parameters are integrated into regulatory studies and may provide valuable functional explanations for toxicology findings.

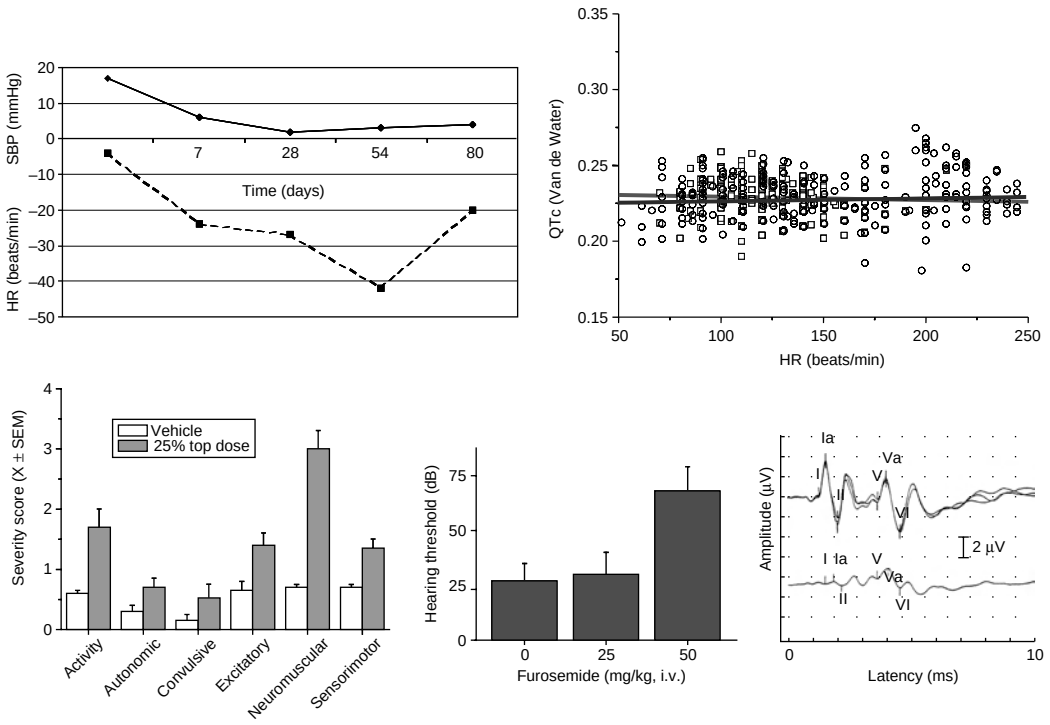


FIGURE 13.4

Examples illustrating the measurement of safety pharmacology endpoints in toxicology studies.

Top left panel — Monitoring of systolic blood pressure (SBP) and heart rate (HR) in rats over a 3-month period following once daily treatment. Values are vehicle subtracted mean absolute changes in SBP and HR. SBP was measured using the tail cuff method and HR was derived from pulsatile signal of BP. *Top right panel* — Drug/vehicle effect on Van de Water corrected QT interval in dogs treated once daily for up to 1 year. Data are from 177 dogs. Baseline (predose) values or vehicle data are presented as open squares, whereas postdose or drug-treated data are plotted as open circles. The regression lines were not statistically, significantly different from each other and an the slopes of the regression lines were not different from zero. *Bottom right panel* — Brainstem auditory evoked responses (BAER) induced by auditory stimuli via an earpiece were recorded from three conscious restrained beagles dogs using subdermal needle electrodes via a medelec Sapphire IIA system. Repeated determinations were made of background auditory threshold in decibels (dB) and the amplitudes and latencies of waveforms I to V at 80 dB. Furosemide (25 or 50 mg/kg) was then administered intravenously and the effects on these parameters assessed over a 60-min period. Preliminary results demonstrated that Frusemide at 50 mg/kg but not 25 mg/kg increased the hearing threshold in dogs by approximately 40 dB (left-hand side). In addition, the amplitudes of waves I to V were markedly reduced (right-hand side). These effects were transient returning to normal by 6-min post-dose. *Bottom left panel* — Neurobehavioral assessment using a functional observational battery and an automated assessment of motor activity. Dose- and time-related effects of seven prototypic chemicals following both single and 4-week repeated exposures were tested in four laboratories in the United States and four in Europe. The results indicated that neurotoxicants could be detected and characterized, despite some differences on specific endpoints. This study also provides extensive data regarding the use of neurobehavioral screening methods over a range of laboratory conditions as well as the reliability, sensitivity, and robustness of the tests to detect neurotoxic potential of chemicals.[197] Domain scores represent the neurotoxicity of acrylamide after 4 weeks of dosing (5 days/week). The scores for each functional domain are averaged across all laboratories, and data are presented as means ±SEM. For clarity, only the vehicle and the 25% of the top dose groups are shown.

13.5 Predictive Value of Non-Clinical Safety Pharmacology Testing to Humans

It is difficult to answer the question "Are non-clinical safety pharmacology tests predictive of side effects in humans?" Primarily because hard evidence is not readily available in the public domain, if an NCE has an effect in a non-clinical test, there may be limited information in the public domain, as the result may have precluded clinical development and therefore no apparent value in communicating this information. If there has been no effect in a non-clinical test, and likewise no effect in the corresponding variable in humans, these negative data may have been deemed not to be of publishable interest. What we are then left with are the high profile examples of side effects in humans that are apparently not detected non-clinically. One publication attempted to explore the predictive values of safety pharmacology assays to humans.[6] Some significant correlations were reported. For example, a decreased locomotor activity in rodents was positively correlated with dizziness and sleepiness in humans; a decreased intestinal transit in rodents was correlated with constipation and 'anorexia in humans; and a decreased urinary and sodium excretion in the rat was correlated with edema in humans. Rather more bizarrely, the findings of analgesia, decreased body temperature, and anticonvulsive activity in rodents were each correlated with "thirst" in humans. This indicates the limitations of such surveys.

There are, however, numerous examples of drugs that cause adverse effects in humans, which would be detectable in safety pharmacology studies conducted as per ICH S7A and S7B guidances.

Some notable examples of individual drugs showing untoward effects in non-clinical studies that are correlated in a quantitative sense with adverse effects in humans have been reported.[2] For example: (1) the sedative effects of clonidine in various animal species and man; (2) the propensity of cisapride to prolong ventricular repolarization; (3) the respiratory depressant effects of morphine; (4) the nephrotoxic effect of cyclosporine; and (5) the GI effects of erythromycin. These examples illustrated the very good agreement of effects across all species tested and across a narrow range of doses/concentrations. Over the last few years, data have been generated to assess the value of non-clinical tests to predict the potential of NCEs to prolong the QT interval of the ECG and ultimately the proarrhythmic potential of these drugs. The published data converged in that an integrated risk assessment (see below) based upon data on the potency against hERG, an *in vivo* repolarization assay and if necessary an *in vitro* repolarization assay are in a qualitative sense predictive of the clinical outcome.[198–201] These data have been further supported by publications suggesting that a 30-fold margin between the highest free plasma concentration of a drug in clinical use (C_{max}) and the concentration inhibited by 50% of the hERG current (IC_{50}) could be adequate to insure an acceptable degree of safety from arrhythmogenesis with a low risk of obtaining false positives.[38,39,202]

In the same way as in efficacy and kinetic models, there are, however, various reasons why non-clinical safety pharmacology assays may not predict human adverse effects: (1) species differences in the expression or functionality of the molecular target mediating the adverse effects; (2) differences in pharmacokinetic properties between test species and man; (3) sensitivity of the test system (e.g., observations of a qualitative nature should be followed-up with specific quantitative assessment); (4) poor optimization of the test conditions (the baseline level has to be set correctly to detect drug-induced changes); (5) study designs that are statistically underpowered; (6) inappropriate timing of functional measurements in relation to the time of maximal effect (i.e., T_{max}); (7) delayed effects (safety pharmacology studies generally involved a single-dose administration with time points covering the pharmacokinetic profile of the parent drug); (8) difficulty of detection in animals (adverse effects such as arrhythmia, headache, disorientation, and hallucinations are

quite a challenge to detect in safety pharmacology studies); and (9) assessment of a suboptimal surrogate endpoint that predicts with some degree of confidence the clinical outcome (e.g., QT/QTc interval prolongation as a surrogate of TdP).

13.6 Integrated Risk Assessment

The integrated risk assessment is the stepwise and holistic evaluation of non-clinical study results in conjunction with any other relevant information and should be scientifically based and individualized for an NCE. Such an assessment can contribute to the design of clinical investigations and the interpretation of their findings.

Risk assessment in terms of protecting Phase I clinical trial participants is relatively straightforward, as it does not take into account any consideration of the therapeutic target (unless when the Phase I trials include patients) or the degree of unmet medical need. Therefore, the assessment of the safety pharmacology data has to take into consideration the severity of the outcome in any given safety pharmacology test (see below— bearing in mind the sensitivity and specificity of the assays), and the plasma concentration at which it occurred relative to the expected exposure in the trial. Depending on the stage of drug development, the integrated risk assessment should consider contribution of metabolites as well as metabolic differences between humans and animals.

In terms of early risk assessment of NCE viability, the situation is more complex. In an attempt to simplify and standardize how safety pharmacology data can contribute to early risk assessment of project viability, Redfern et al.[2] proposed a matrix type approach described below. This requires a grading process; each of the factors in the risk assessment can be graded into, for example, three categories — low, medium, and high. Starting with the safety pharmacology tests themselves, they can be categorized as follows: (1) minor — predictive of nonserious, reversible side effects (e.g., certain GI or renal effects); (2) moderate — predictive of impairment of quality of life (e.g., sedation, motor coordination); (3) major — predictive of potentially life-threatening effects (e.g., prolonged the QT interval, pronounced hypotension, or bronchoconstriction). The next step consists in grading the therapeutic target according to disease severity: (1) minor/moderate disease (e.g., eczema, rhinitis, and Raynaud's syndrome); (2) debilitating disease (e.g., asthma, epilepsy, Parkinson's disease, stroke, and angina); and (3) life-threatening disease (e.g., cancer, AIDS, and myocardial infarction). The third component that can be considered is the existing therapy; the potential new drug must be anticipated to be superior to existing therapy. Therefore, the existing therapy cannot be classified as excellent; instead it can be rated as (1) good; (2) partially effective with side effects; and (3) poor/inexistent. Once collected, the set of information can be put together in a matrix that also takes into account the dose level (or concentration) at which the effects were observed in the safety pharmacology test, in comparison to the expected clinical exposure (total or free plasma concentration), as shown in Table 13.11. Effects well to the left of the line of crosses are acceptable without further debate, those well to the right are unacceptable, and those on or near the line require further discussion. Hypothetical examples are presented to illustrate the usefulness of such matrix. In the first example, an NCE targeted at Raynaud's syndrome — a minor/moderate; disease[2] under classification, for which existing therapy is poor. The NCE is found to block the hERG channel and prolong the QT interval at a relatively low multiple (e.g., ~30-fold) above the expected therapeutic plasma concentration (Figure 13.5 as a schematic illustration). According to the matrix analysis, the decision is either to discontinue the progression of the NCE into development or to accept embarking on an extensive and expensive clinical program in compliance with the ICH E14 guidance.[53] At

TABLE 13.11

Safety Pharmacology Integrated Risk Assessment Matrix

		Severity and Acceptability of Safety Pharmacology Outcome								
		Minor			Moderate			Major		
Therapeutic Target	**Existing Therapy**	**100×**	**10×**	**1×**	**100×**	**10×**	**1×**	**100×**	**10×**	**1×**
Minor disease	Good	X								
	Partially effective/ side effects		X							
	Poor/none			X						
Debilitating disease	Good				X					
	Partially effective/ side effects					X				
	Poor/none						X			
Life-threatening disease	Good							X		
	Partially effective/ side effects								X	
	Poor/none									X

Note: Outcomes to the left of the line of crosses are acceptable, those to the right are unacceptable, whereas outcomes on or near the line require further discussion and possibly further investigations. 100X, 10X, 1X represents 100, 10 or 1 fold the therapeutic dose.

Source: Adapted from Redfern, W.S. et al., *Fundam. Clin. Pharmacol.*, 16, 161–173, 2002.

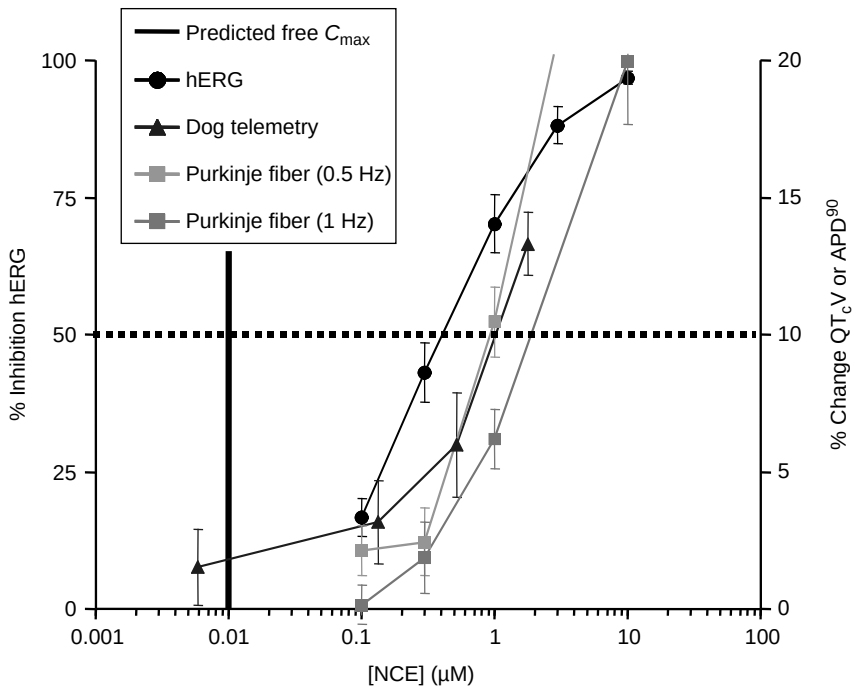


FIGURE 13.5

Schematic illustration of an integrated risk assessment aimed at assessing the liability for an NCE to prolong the QT interval in man. APD = Action Potential Duration; NCE = New Chemical Entity; QT_cV = Van de Water corrected QT interval duration.

the other end of the spectrum, an NCE is targeted at leukemia, where the existing therapy is partially effective and has marked side effects. The NCE was found to increase intestinal motility at the therapeutic dose; this would not be a major issue to prevent progression into clinical development. However, this could become an issue if the NCE eventually reaches the market and has to compete with a new rival drug of similar efficacy but lacking this side effect. Other factors should be considered such as (1) the target population, for example, cognitive impairment as a side effect may be more problematic in elderly and pediatric patients; and (2) the ultimate project objective, for example, QT interval prolongation as a side effect may be manageable and acceptable for an NCE aimed at demonstrating proof of mechanism or proof of principle. Overall, the evidence of risk, as part of an integrated risk assessment, can support the planning and interpretation of subsequent clinical studies.

13.7 Current and Future Challenges

The future of safety pharmacology will depend, in part, upon the scientific and technological advances and regulatory challenges that envelop pharmaceutical development.

13.7.1 Science and Technology

With advances in molecular biology and biotechnology, which allows for the identification of new clinical molecular targets, newer pharmaceutical agents are being identified that act at these novel molecular sites in an attempt to ameliorate the disease condition. Moreover, new therapeutic approaches are being developed (e.g., gene therapy and biotech products) that present new challenges to the safety pharmacologist. Inherent in the novelty of new targets and new approaches is the risk of unwanted effects that may or may not be detected based on current scientific knowledge and techniques. The scientific challenge facing safety pharmacology is to keep pace, to adapt, and to incorporate new technologies in the evaluation of new drugs in non-clinical models and identifying the effects that pose a risk to human volunteers and patients. Recent examples include safety embracement of modern electrophysiological techniques to evaluate the effects of NCEs on ionic components of the cardiac action potential,[39,83] and telemetry techniques to permit the chronic monitoring of physiological functions in unstressed animals.[186,203–205] Efforts continue to construct databases that relate to the predictive value of non-clinical assays to man either through retrospective analysis[4,6,39] or through purposely designed studies.[198–201] As an example, non-clinical safety studies are typically conducted in normal, healthy, young adult, or adult animals. However, these tests may not appropriately detect specific responses in humans at other ages (e.g., neonates, adolescents, and geriatrics) or those with underlying chronic diseases (e.g., heart failure, renal failure, and Type II diabetes), conditions which may alter the pharmacodynamic response to an NCE. In some cases, animal models that overexpress or are deficient in the unique molecular targets, or are otherwise manipulated to model the human pathophysiological conditions, may provide additional focus and sensitivity to detect and interpret the potential unwanted effects of NCEs in terms of human risk (e.g., assessment of the proarrhythmic potential of NCEs).[40,41] Translating understanding of normal and perturbed physiological function into the theater of non-clinical safety assessment is, therefore, a major challenge. The challenge is to identify non-clinical models that reflect the overall human pathological condition and to incorporate these disease models along with traditional safety models/assays

into safety pharmacology paradigms to produce integrated and more accurate assessment of possible human risk. For example, the detection of NCE-induced acute renal failure is poorly predicted in normal healthy animals, whereas the sensitivity is increased in sensitized animals that have reduced glomerular filtration reserve. Another emerging approach is the use of adult stem cells to evaluate cardiotoxicity.[206]

Adult stem cells-derived cardiomyocyte have been successfully obtained. These cells present characteristics of differentiated ventricular cardiomyocyte, including a typical action potential. The ability to expand human adult stem cells *in vitro* provides unlimited potential for producing quantities of cardiomyocyte –progenitors, which could be employed to assess the effects of NCEs on cardiac ion channels, electrical activity, and contractility.[206] The challenge posed by the introduction of new techniques and technologies in formulating a risk assessment is to improve and enhance the safe progression of an NCE to the marketplace, while preventing unnecessary delays (or discontinuances), based on non-clinical findings that are not relevant or interpretable in terms of clinical response or human risk. Ultimately, increase in scientific knowledge and development of new technologies should lead to the development of new, robust, and predictive assays/models for ADRs that are currently poorly predicted such as nausea, headache, and arrhythmia. Below are the two examples of promising future areas for development of safety pharmacology that further illustrate the challenges in front of us.

13.7.1.1 In Vivo *Approach*

The zebrafish (*Danio rerio*) is a small tropical fish, which develops *ex utero* and is optically transparent during early development. The larvae (1 to 4 mm long) can live for up to ~7 days in multiwell plates (up to 384 wells) because they are supported by nutrients in their yolk sac. Few days after fertilization, zebrafish develops discrete organs and tissues, including brain, heart, liver, pancreas, kidney, intestines, bones, muscles, nervous systems, and sensory organs. These organs and tissues have been demonstrated to present similarities to their mammalian counterparts at the anatomical, physiological, and molecular levels. Compounds can be added directly to the wells, and only small (<mg) quantities are required for screening studies. This methodology is gaining acceptance in toxicity testing and drug discovery, as means of frontloading *in vivo* studies.[207–210] Although, in most instance, it is an *in vivo* phenotype-based (i.e., black box) method of screening, some models comply with a conventional strategy of a target-based (i.e., rational design/rational evaluation) approach. The assays are amenable to medium/high throughput screen, chronic dosing (i.e., up to several days), and combination therapies. Models have been developed to assess drug effects primarily on cardiovascular, nervous (including sensory functions), gastrointestinal, and immune systems. A non-exhaustive list of models amenable to assess drug effects on physiological functions that may have relevance to safety pharmacology is presented in Table 13.12. The methods used are either directly transferred from mammalians and adapted to the zebrafish (e.g., startle reflex) or use technologies amenable to the characteristics of the zebrafish (e.g., digital video imaging recording to determine heart rate). Although the preliminary published observations are encouraging, each model would require thorough, extensive validation using a wide range of compounds to insure its reliability, reproducibility, and predictive value to man. Although the zebrafish larvae appear as an attractive potential non-clinical model, several limitations should be recognized and addressed prior to get confidence in introducing this *in vivo* model for safety pharmacology testing. The areas for future development are as follows: (1) robust characterization (physiological and morphological) and pharmacological validations of each of the functional endpoints; (2) reproducibility over time and across laboratories; and (3) understanding and characterization of the kinetics properties, exposure to the NCEs and metabolism capabilities of the zebrafish larvae.

TABLE 13.12

Non-Exhaustive List of Potential Zebrafish Models Amenable to Assess Drug Effects on Physiological Functions Relevant to Safety Pharmacology

Physiological Function	Endpoint	Method/Technique	References
Gastrointestinal	Peristalsis (number and amplitude of the contractions)	Video microscopy of fluorescent diets	Farber et al.;[211] Wallace and Pack[212]
Pro-/anticonvulsive potential	Stereotyped behavioral changes and electrographic discharges	Chemically induced seizures	Baraban et al.[213]
Behavioral (e.g., sleep, aggression and swim patterns, locomotor skills)	Stereotyped behavioral changes		Pasini and Wilkinson;[214] Gerlai et al.;[215] Gerlai;[216] Zhdanova et al.[217]
Visual	Stereotyped behavioral response to a sudden decrease in illumination	Optokinetic/optomotor responses	Kimmel et al.[218]
Auditory	Stereotyped behavioral response to a sudden sound	Acoustic startle response by video tracking	Bang et al.;[219] Zeddies and Fay;[220] Whitfield[221]
Vascular	Vasoactive responses — Measurement of changes in vessel diameter	Digital motion analysis	Fritsche et al.[222]
Cardiovascular	Ventricular and blood pressures	Servonull system	Hu et al.[223]
Cardiovascular	Bradycardia and arrhythmia (block of atrio-ventricular conduction) used as surrogate of drug-induced QT prolongation		Milan et al.;[224] Ulrike et al.[225]
Hematology and coagulation	Red blood cell concentration, edema, hemorrhage, bleeding time, aggregation	Digital video imaging (in presence or absence of fluorescent beads)	Schwerte et al.;[226] Grosser et al.[227]
Immune	T-cell-mediated immunosuppression	Transgenic zebrafish associated with fluorescent imaging	Langenua et al.[228]
Pharmacological profiling	Muscarinic acetylcholine M2 receptor	Binding and functional response (bradycardia)	Hsieh and Liao[229]

13.7.1.2 In Silico *Approach*

There is a considerable interest in computational models to predict the safety of NCEs in the drug discovery and development phases. Insight into the safety pharmacological potential of a scaffold or series of structures early in the drug discovery process could help the medicinal chemists to prioritize particular scaffolds or hits or alternatively can contribute halting the discovery process for a given research project. The main safety

pharmacology area that has been explored in terms of computational evaluation and prediction is cardiotoxicity. Inhibition of the hERG channel and its link to QT prolongation and TdP have been the topic of several computational modeling papers.[230] Here the challenge is that both potent and weak blockade of the hERG channel can increase the risk of compound series showing unwanted cardiac effects *in vivo*. Indeed, in many therapeutic areas, such as in antibiotics, analgesics, and oncology drugs, therapeutic concentrations can well be in the micromolar range, hence even weak to moderate hERG block would need to be accurately predicted.

Another challenge is that the crystallography structure of the hERG channel is not yet available, therefore some models have been constructed from the crystal structure of the bacterial MthK potassium channel, which obviously present some limitations. Several papers have rationalized compound interactions with hERG channel using both small and medium size measured or literature datasets and have proposed purported predictive but separate models.[231–233] The diversity of these models places the medicinal chemist in front of a dilemma. Which of those models are best suited for particular scaffolds? Would any of the available models be predictive in the chemical space explored by a project team? Therefore, computational tools that suggest different options for chemists could have some benefits here. Toward that goal, Zolotoy et al.[234] describe the physicochemical determinants for a potent hERG inhibition and give suggestions on structural features that reduce the potency of hERG blockade. There are now several models, filters, and decision trees based solely on physicochemical descriptors of hERG blockers. There may also be value for hERG model that would accurately predict molecules that have a low potency at hERG (e.g., $IC_{50}<10\ \mu M$). Such model with a low false-positive rate would identify only a small portion of the hERG blockers, but could complement inadequate experimental methods. The value comes from the fact that (1) many of the high-throughput experimental methods for hERG blockade underestimate the potency of hERG blockade as measured by conventional electrophysiology techniques[235] and (2) molecules in early discovery projects often have poor physicochemical properties, especially poor solubility. Therefore, such molecules and chemical series can escape early detection of their inhibition of the hERG channel but reveal their hERG characteristics in the early phases of drug discovery where more soluble members of the chemical series are available. A computational trigger indicating high likelihood of hERG block could, therefore, prompt a team to follow-up with a relevant approach. Although it is recognized that hERG is probably a key contributor to drug-induced QT prolongation and TdP, some other molecular mechanisms are likely to be involved including other ion channels. Align to the fact that experimental and clinical possibilities for studying cardiac arrhythmias in human ventricular myocardium are very limited, the use of alternative methods such as computer simulation is of great interest.[236,237] It is worth mentioning that mathematical models of the human ventricular action potential are being developed.[238] Such computer modeling of the heart from genes to cells to the whole organ is becoming a reality that with no doubt will significantly impact on the way safety pharmacology evolves in the future.[237,239] It is important to mention that molecular modeling in safety pharmacology is not restricted to the cardiovascular system, valuable models predicting blood–brain barrier partitioning of NCEs have already been described.[240] Ultimately, such models should be key in defining the overall CNS program that a given NCE should go through.

13.7.2 Regulatory Requirements

The future of safety pharmacology is intertwined with international regulatory guidance. The ICH S7B and E14 guidances have been finalized and are due for implementation in November 2005.[25,53] An emerging draft guidance from the European Medicines Agency for the Evaluation of Medicines for Human Use on the non-clinical investigation of the

dependence potential of medicinal products has been released for consultation and once finalized will influence the overall approach to safety pharmacology especially for CNS-targeted NCEs.[60] Safety pharmacology is also considered an important component to newly emerging regulatory guidance from the U.S. FDA such as the safety evaluation of pediatric drug products and non-clinical studies for development of pharmaceutical excipients. The discipline is considered integral to the evolving regulatory strategies for safety that accelerates the introduction of NCE into clinical phases (cf., Position paper on non-clinical safety studies to support clinical trials with a single microdose[241]) and the U.S. FDA, screening investigational new drug application.[45,242] It is worth elaborating on this topic since it is likely to influence the approach to safety pharmacology. In its March 2004 Critical Path Report, the FDA explained that to reduce the time and resources expended during the early drug development on NCEs that are unlikely to succeed, tools are needed to distinguish earlier in the process those candidates that hold promise from those that do not.[243] The FDA suggests that limited exploratory IND investigations in humans can be initiated with less, or different, non-clinical support that is required for traditional IND studies because exploratory IND studies present fewer potential risks than do traditional Phase I studies that look for dose-limiting toxicities.[45] The draft guidance describes some early Phase I exploratory approaches that are consistent with regulatory requirements, but that will enable sponsors to move ahead more efficiently with the development of promising NCEs while maintaining needed human subject protections. Exploratory IND studies, which usually involve very limited human exposure and have no therapeutic intent, can serve a number of useful goals such as to (1) gain an understanding of the relationship between a specific mechanism of action and the treatment of a disease; (2) provide important information on pharmacokinetics properties of an NCE; (3) select the most promising lead product from a group of NCEs designed to interact with a particular therapeutic target in humans; and (4) explore the characteristics of the bio-distribution of a product using various imaging technologies. The non-clinical safety evaluation recommended for an exploratory IND application is more limited than for a traditional IND application[24] and may be tailored to the intended clinical study design and objective. For example, if the aim is to administer a microdose (i.e., 1/100th of the pharmacological dose and $<100\ \mu g$), safety pharmacology endpoints may be incorporated into the single-dose toxicity studies (e.g., evaluation of respiratory and central nervous systems using clinical observations or FOB/Irwin's tests usually in rodent species) although we would strongly recommend an appropriate quantitative assessment of these vital organ functions. Furthermore, a comprehensive *in vitro* pharmacological profiling, aimed at determining the affinity and activity of the NCE at molecular targets known to be associated with unwanted activities, should be recommended (e.g., hERG, α-adrenoceptor). If the clinical trial is designed to study pharmacological effects or to investigate the mechanism(s) of action, then the effects of the NCE should be evaluated in safety pharmacology studies. Evaluation of the central nervous and respiratory systems can be performed as part of the rodent toxicology studies while evaluation of the cardiovascular system can be assessed in the nonrodent species, generally the dog. The replacement of stand-alone safety pharmacology studies by the inclusion of safety pharmacology endpoints into toxicology studies has clearly some benefit but also some potential drawbacks that one has to be aware, accept, and recognize the impact on quality and therefore the value of the data (see Section 13.4.3.3).

The introduction of pharmaceuticals into the environment is gaining growing attention of both regulators and pharmaceutical industry.[244–246] While this is not currently the subject of any international environmental guideline, the use of organ function endpoints may become an important component in bridging safety data collected from mammalian vertebrates (including humans) to aquatic species for purposes of the identification of relevant target and organ functions and the design of specific environmental toxicology studies.

13.7.3 Training and Education

Safety pharmacology also faces significant challenges of attracting, training, and certifying investigators in integrative approaches to physiology and pharmacology to insure the development of its promising new future.[247,248] The paucity of training in integrative biomedical sciences has had detrimental long-lasting effects such as (1) an impact on the development of intact animal models of human function and disease; (2) an impact on skills to conceptualize biomedical hypothesis and experiments at the level of the intact animal; and (3) an impact on the process of non-clinical and clinical drug discovery and development.[248]

13.8 Conclusion

Safety pharmacology is a rapidly growing and evolving discipline that is facing significant challenges on scientific, technological, regulatory, and human fronts. Thus, the answer to the future challenges of safety pharmacology will be contained within the vision of its current and future leaders, the issues and concerns that they face, and the solutions to the important challenges that they generate.

Acknowledgments

The authors wish to thank members of the Safety Pharmacology Department, Safety Assessment UK for their contribution over the years enabling this chapter to be written. Our thanks go to Najah Abi-Gerges, Duncan Armstrong, Matthew Bridgeland-Taylor, Joanne Bowes, Alison Easter, Dave Hamre, Alex Harmer, Chris Lawrence, Roope Mannikko, Nick McMahon, Jackie Moors, Karen Philp, Chris Pollard, Helen Prior, Will Redfern, Mike Rolf, Jason Schofield, Ben Small, Nicola Smith, Sharon Storey, Isobel Strang, and Ann Woods.

References

1a. Kola, I. and Landis, J., Can the pharmaceutical industry reduce attrition rates?, *Nat. Rev. Drug Discov.*, 3, 711–715, 2004.
1b. Lasser, K.E., Allen, P.D., Woolhandler, S.J., Himmelstein, D.U., Wolfe, S.M., and Bor, D.H., Timing of new black box warnings and withdrawals for prescription medications, *JAMA,* 287, 2215–2220, 2002.
1c. Kennedy, T., Managing the drug discovery/development interface, *Drug Discov. Dev.,* 2, 436–444, 1997.
2. Redfern, W.S., Wakefield, I.D., Prior, H., Hammond, T.G., and Valentin, J.P., Safety pharmacology — A progressive approach, *Fundam. Clin. Pharmacol.*, 16, 161–173, 2002.
3. Sibille, M., Deigat, N., Janin, A., Kirkesseli, S., and Durand, D.V., Adverse events in phase-I studies: a report in 1015 healthy volunteers, *Eur. J. Clin. Pharmacol.*, 54, 13–20, 1998.
4. Olson, H., Betton, G., Robinson, D., Thomas, K., Monro, A., Kolaja, G., Lilly, P., Sanders, J., Sipes, G., Bracken, W., Dorato, M., Van Deun, K., Smith, P., Berger, B., and Heller, A., Concordance of the toxicity of pharmaceuticals in humans and in animals, *Regul. Toxicol. Pharmacol.*, 32, 56–67, 2000.

5. Greaves, P., Williams, A., and Eve, M., First dose of potential new medicines to humans: how do animals help?, *Nat. Rev. Drug Discov.*, 3, 226–236, 2004.
6. Igarashi, T., Nakane, S., and Kitagawa, T., Predictability of clinical adverse reactions of drugs by general pharmacology studies, *J. Toxicol. Sci.*, 20, 77–92, 1995.
7. Stephens, M.D.B., in *Stephens' Detection of New Adverse Drug Reactions*, 5th ed., Talbot, J. and Waller, P., Eds., 2004, pp. 1–91.
8. Lazarou, J., Pomeranz, B.H., and Corey, P.N., Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies, *JAMA*, 279, 1200–1205, 1998.
9. McKenney, J.M., Pharmacologic options for aggressive low-density lipoprotein cholesterol lowering: benefits versus risks, *Am. J. Cardiol.*, 96, 60E–66E, 2005.
10. Fung, M., Thornton, A., Mybeck, K., Wu, J.H., Hornbuckle, K., and Muniz, E., Evaluation of the characteristics of safety withdrawal of prescription drugs from worldwide pharmaceuticals markets — 1960 to 1999, *Drug Inform. J.*, 35, 293–317, 2001.
11. Gad, S.C., Ed., *Safety Pharmacology in Pharmaceutical Development and Approval*, CRC Press, Boca Raton, FL, 2004, pp. 95–108.
12. Kinter, L.B., Gossett, K.A., and Kerns, W.D., Status of safety pharmacology in the pharmaceutical industry — 1993, *Drug Dev. Res.*, 32, 208–216, 1994.
13. Lumley, C.E., General pharmacology, the international regulatory environment, and harmonization guidelines, *Drug Dev. Res.*, 32, 223–232, 1994.
14. Green, M.D., An FDA perspective on general pharmacology studies to determine activity and safety, *Drug Dev. Res.*, 35, 158–160, 1995.
15. Proakis, A.G., Regulatory considerations on the role of general pharmacology studies in the development of therapeutic agents, *Drug Dev. Res.*, 32, 233–236, 1994.
16. Anon., Guidelines for general pharmacology studies–Japanese guidelines for non-clinical studies of drugs manual 1995 (pp. 71–80), Eds., Nippo Yakuji, Ltd.
17. Kinter, L.B. and Valentin, J.P., Safety pharmacology and risk assessment, *Fundam. Clin. Pharmacol.*, 16, 175–182, 2002.
18. Bass, A.S., Kinter, L.B., and Williams, P., Origins, practices and future of safety pharmacology, *J. Pharmacol. Toxicol. Methods*, 49, 145–151, 2004.
19. Bass, A.S. and Williams, P.D., Status of international regulatory guidelines on safety pharmacology, in *Safety Pharmacology*, Williams, P.D. and Bass, A.S., Eds., XXX, pp. 9–20.
20. Bowes, J., Rolf, M., and Valentin, J.P., Pharmacological profiling: the early assessment of pharmacological properties, in *The Process of New Drug Discovery and Development*, Smith, C.G., Ed., J. O'Donnell, 2006, chap. 6, this volume.
21. Kurata, M., Kanai, K., Mizuguchi, K., Yoshida, M., Nakamura, K., Furuga, Y., Kinoue, A., Totsuka, K., and Igarashi, S., Trends in safety pharmacology in the US and Europe, *J. Toxicol. Sci.*, 22, 237–248, 1997.
22. Anon., European Medicines Evaluation Agency, Human Medicines Evaluation Unit, Committee for Proprietary Medicinal Products, Points to Consider: The assessment of QT interval prolongation by non-cardiovascular medicinal products, CPMP/986/96, London, December 17, 1997, http://www.emea.eu.int/pdfs/human/swp/098696en.pdf
23. Anon., ICH S6: Preclinical safety evaluation of biotechnology-derived pharmaceuticals, CPMP/ICH/302/95, London, July 16, 1997.
24. Anon., ICH M3: Timing of non-clinical safety studies for the conduct of human clinical trials for pharmaceuticals, CPMP/ICH/286/95, 1997, Modified version released November 2000.
25. Anon., ICHS7A: Safety pharmacology studies for human pharmaceuticals, CPMP/ICH/539/00, London, November 16, 2000, http://www.emea.eu.int/pdfs/human/ich/053900en.pdf
26. Anon., *The GLP Pocket Book: The Good Laboratory Practice Regulations 1999 and Guide to UK GLP Regulations 1999*, MCA Publications, London, pp. 1–75.
27. Anon., Japanese Ministry of Heath and Welfare (MHW) in: Notes on applications for approval to manufacture new drugs, 1975.
28. Anon., Recommendations from the scientific advisory panel subgroups on hepatotoxicity: Hepatotoxicity of health products, 2004.
29. Ackerman, M.J., The long QT syndrome: ion channel diseases of the heart, *Mayo Clin. Proc.*, 73, 250–269, 1998.

30. Anderson, M.E., Al-Khatib, S.M., Roden, D.M., and Califf, R.M., Cardiac repolarization: current knowledge, criticl gaps and new approaches to drug development and patient management, *Am. Heart J.*, 144, 769–781, 2002.
31. De Ponti, F., Poluzzi, E., and Montanaro, N., Organising evidence on QT prolongation and occurrence of Torsades de pointes with non-anti-arrhythmic drugs: a call for consensus, *Eur. J. Clin. Pharmacol.*, 57, 185–209, 2001.
32. Haverkamp, E., Breithart, G., Camm, A.J., Janse, M.J., Rosen, M.R., Antzelevitch, C., Escande, D., Franz, M., Malik, M., Moss, A., and Shah, R., The potential for QT prolongation and proarrhythmia by non-antiarrhythmic drugs: clinical and regulatory implications, *Cardiovasc. Res.*, 47, 219–233, 2000.
33. Fenichel, R.R., Malik, M., Antzelevitch, C., Sanguinetti, M., Roden, D.M., Priori, S.G., Ruskin, J.N., Lipicky, R.J., and Cantilena, L.R., Drug induced torsades de pointes and implication for drug development, *J. Cardiovasc. Electrophysiol.*, 15, 475–495, 2004.
34. Malik, M. and Camm, A.J., Evaluation of drug-induced QT interval prolongation. Implications for drug approval and labeling, *Drug Safety*, 24, 323–351, 2001.
35. Moss, A.J., The QT interval and torsade de pointes, *Drug Safety*, 21, 5–10, 1999.
36. Thomas, S.H.L., Drugs, QT interval abnormalities and ventricular arrhythmias, *Adverse Drug React. Toxicol. Rev.*, 13, 77–102, 1994.
37. Viskin, S., Long QT syndromes and torsade de pointes, *Lancet*, 354, 1625–1633, 1999.
38. Webster, R., Leischmann, D., and Walker, D., Towards a drug concentration effect relationship for QT prolongation and torsades de pointes, *Curr. Opin. Drug Discov. Dev.*, 5, 116–126, 2002.
39. Redfern, W.S., Carlsson, L., Davis, A.S., Lynch, W.G., MacKenzie, I., Palethorpe, S., Siegl, P.K.S., Strang, I., Sullivan, A.T., Wallis, R., Camm, A.J., and Hammond, T.G., Relationship between preclinical cardiac electrophysiology, clinical QT interval prolongation and torsade de pointes for a broad range of drugs: evidence for a provisional safety margin in drug development, *Cardiovasc. Res.*, 58, 32–45, 2003.
40. Valentin, J.P., Hoffmann, P., De Clerck, F., Hammond, T.G., and Hondeghem, L., Review of the predictive value of the Langendorff heart model (Screenit system) in assessing the proarrhythmic potential of drugs, *J. Pharmacol. Toxicol. Methods*, 49, 171–181, 2004.
41. Lawrence, C.L., Pollard, C.E., Hammond, T.G., and Valentin, J.P., Nonclinical proarrhythmia models: predicting Torsades de Pointes, *J. Pharmacol. Toxicol. Methods*, 52, 46–59, 2005.
42. Anon., ICH S7B: The nonclinical evaluation of the potential for delayed ventricular repolarization (QT interval prolongation) by human pharmaceuticals, CPMP/ICH/423/02, London, 25 May 2005, http://www.emea.eu.int/pdfs/human/ich/042302en.pdf
43. Valentin, J.P., Bass, A.S., Atrakchi, A., Olejniczak, K., and Kannosuke, F., Challenges and lessons learned since implementation of the safety pharmacology guidance ICH S7A, *J. Pharmacol. Toxicol. Methods,* 52, 22–29, 2005.
44. Anon., Guideline on the evaluation of control samples in nonclinical safety studies: checking for contamination with the test substance, CPMP/SWP/1094/04, London, 17 March 2005.
45. Anon., Guidance for industry, investigators and reviewers: Exploratory IND studies, 2005, http://www.fda.gov/cder/guidance/6384dft.pdf
46. Zbinden, G., Neglect of function and obsession with structure in toxicity testing, *Proceedings of the 9th International Congress of Pharmacology*, 1984, Vol. 1, pp. 43–49.
47. Dressman, J.B., Comparison of canine and human gastrointestinal physiology, *Pharm. Res.*, 3, 123–131, 1986.
48. Schein, P.S., Davis, R.D., Carter, S., Newman, J., Schein, D.R., and Rall, D.P., The evaluation of anticancer drugs in dogs and monkeys for the prediction of qualitative toxicities in man, *Clin. Pharmacol. Ther.*, 11, 3–40, 1970.
49. Abi-Gerges, N., Philp, K., Pollard, C., Wakefield, I., Hammond, T., and Valentin, J.P., Sex differences in ventricular repolarization: from cardiac electrophysiology to Torsades de Pointes, *Fund. Clin. Pharmacol.*, 18, 139–151, 2004.
50. Festing, M.F.W., Overend, P., Gaines Das, R., Cortina Borja, M., and Berdoy, M., *Laboratory Animal Handbooks No. 14, The Design of Animal Experiments*, Published for Laboratory Animals Ltd by Royal Society of Medicine Press Ltd, London, 2003, pp. 71–83.

51. Chiang, A.Y., Smith, W.C., Main, B.W., and Sarazan, R.D., Statistical power analysis for hemodynamic cardiovascular safety pharmacology studies in beagle dogs, *J. Pharmacol. Toxicol. Methods*, 50, 121–130, 2004.
52. Baillie, T.A., Cayen, M.N., Fouda, H., Gerson, R.J., Green, J.D., Grossman, S.J., Klunk, L.J., LeBlanc, B., Perkins, D.G., and Shipley, L.A., Drug metabolites in safety testing, *Toxicol. Appl. Pharmacol.*, 182, 188–196, 2002.
53. Anon., ICH E14: The clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for non-antiarrhythmic drugs, CPMP/ICH/2/04, London, 25 May 2005, http:// www.emea.eu.int/pdfs/human/ich/000204en.pdf
54. Roden, D.M., Drug induced prolongation of the QT interval, *N. Engl. J. Med.*, 350, 1013–1022, 2004.
55. Shah, R.R. and Hondeghem, L.M., Refining detection of drug-induced proarrhythmia: QT interval and TRIaD, *Heart Rhythm*, 2, 758–772, 2005.
56. Belardinelli, L., Antzelevitch, C., and Vos, M.A., Assessing predictors of drug-induced torsade de pointes, *Trends Pharmacol. Sci.*, 24, 619–625, 2003.
57. Straus, S.M.J.M., Sturkenboom, M.C.J.M., Bleumink, G.S., Dieleman, J.P., van der Lei, J., de Graeff, P.A., Kingma, J.H., and Stricker, B.H. Ch., Non-cardiac QTc-prolonging drugs and the risk of sudden cardiac death, *Eur. Heart J.*, 26, 2007–2012, 2005.
58. Volmar, K.E. and Hutchins, G.M., Aortic and mitral fenfluramine-phentermine valvulopathy in 64 patients treated with anorectic agents, *Arch. Pathol. Lab. Med.*, 125, 1555–1561, 2001.
59. Edwards, I.R., What are the real lessons from Vioxx?, *Drug Safety*, 28, 651–658, 2005.
60. Anon., Draft guideline on the non-clinical investigation of the dependence potential of medicinal products, EMEA/CHMP/SWP/94227/2004, London, April 21, 2005.
61. Tang, W., Kang, J., Wu, X., Rampe, D., Wang, L., Shen, H., Li, Z., Dunnington, D., and Garyantes, T., Development and evaluation of high throughput functional assay methods for hERG potassium channel, *J. Biomol. Screen.*, 6, 325–331, 2001.
62. Rezazadeh, S., Hesketh, J.C., and Fedida, D., Rb+ flux through hERG channels affects the potency of channel blocking drugs: correlation with data obtained using a high-throughput Rb+ efflux assay, *J. Biomol. Screen.*, 9, 588–597, 2004.
63. Cheng, C.S., Alderman, D., Kwash, J., Dessaint, J., Patel, R., Lescoe, M.K., Kimrade, M.B., and Yu, W., A high throughput hERG potassium channel function assay: an old assay with a new look, *Drug Dev. Ind. Pharm.*, 28, 177–191, 2002.
64. Sorota, S., Zhang, X.S., Margulis, M., Tucker, K., and Priestley, T., Characterization of a hERG screen using the IonWorks HT: comparison to a hERG rubidium efflux screen, *Assay Drug Dev. Technol.*, 3, 45–57, 2005.
65. Jost, N., Virag, L., Bitay, M., Takacs, J., Lengyel, C., Biliczki, P., Nagy, Z., Bogats, G., Lathrop, D.A., Papp, J.G., and Varro, A., Restricting excessive cardiac action potential and QT prolongation: a vital role for IKs in human ventricular muscle, *Circulation*, 112, 1392–1399, 2005.
66. Liu, D.W. and Antzelevitch, C., Characteristics of the delayed rectifier current (IKr and IKs) in canine ventricular epicardial, midmyocardial, and endocardial myocytes. A weaker IKs contributes to the longer action potential of the M cell, *Circ. Res.*, 76, 351–365, 1995.
67. Jurkiewicz, N.K. and Sanguinetti, M.C., Rate-dependent prolongation of cardiac action potentials by a methanesulfonanilide class III antiarrhythmic agent. Specific block of rapidly activating delayed rectifier K+ current by dofetilide, *Circ. Res.*, 72, 75–83, 1993.
68. Li, G.R., Feng, J., Yue, L., and Carrier, M., Transmural heterogeneity of action potentials and Ito1 in myocytes isolated from the human right ventricle, *Am. J. Physiol.*, 275, H369–H377, 1998.
69. Yang, T. and Roden, D.M., Regulation of sodium current development in cultured atrial tumor myocytes (AT-1 cells), *Am. J. Physiol.*, 271, H541–H547, 1996.
70. Banyasz, T., Fulop, L., Magyar, J., Szentandrassy, N., Varro, A., and Nanasi, P.P., Endocardial versus epicardial differences in L-type calcium current in canine ventricular myocytes studied by action potential voltage clamp, *Cardiovasc. Res.*, 58, 66–75, 2003.
71. Xia, M., Salata, J.J., Figueroa, D.J., Lawlor, A.M., Liang, H.A., Liu, Y., and Connolly, T.M., Functional expression of L- and T-type Ca2+ channels in murine HL-1 cells, *J. Mol. Cell. Cardiol.*, 36, 111–119, 2004.

72. Davie, C., Valentin, J.P., Pollard, C., Standen, N., Mitcheson, J., Alexander, P., and Thong, B., Comparative pharmacology of guinea pig cardiac myocyte and cloned hERG (I(Kr)) channel, *J. Cardiovasc. Electrophysiol.*, 15, 1302–1309, 2004.
73. Suetake, I., Takisawa, H., and Nakamura, T., Related contractile activity and fluorescence changes in fluo-3-loaded isolated ventricular myocytes, *Jpn. J. Physiol.*, 42, 815–821, 1992.
74. Cordeiro, J.M., Greene, L., Heilmann, C., Antzelevitch, D., and Antzelevitch, C., Transmural heterogeneity of calcium activity and mechanical function in the canine left ventricle, *Am. J. Physiol. Heart Circ. Physiol.*, 286, H1471–H1479, 2004.
75. Graham, M.D., Bru-Mercier, G., Hopkins, P.M., and Harrison, S.M., Transient and sustained changes in myofilament sensitivity to Ca2+ contribute to the inotropic effects of sevoflurane in rat ventricle, *Br. J. Anaesth.*, 94, 279–286, 2005.
76. Hamilton, D.L., Boyett, M.R., Harrison, S.M., Davies, L.A., and Hopkins, P.M., The concentration-dependent effects of propofol on rat ventricular myocytes, *Anesth. Analg.*, 91, 276–282, 2000.
77. Witchel, H.J., Milnes, J.T., Mitcheson, J.S., and Hancox, J.C., Troubleshooting problems with in vitro screening of drugs for QT interval prolongation using HERG K+ channels expressed in mammalian cell lines and *Xenopus* oocytes, *J. Pharmacol. Toxicol. Methods*, 48, 65–80, 2002.
78. Zou, A., Curran, M.E., Keating, M.T., and Sanguinetti, M.C., Single HERG delayed rectifier K+ channels expressed in *Xenopus* oocytes, *Am. J. Physiol.*, 272, H1309–H1314, 1997.
79. Cavero, I., Mestre, M., Guillon, J.M., and Crumb, W., Drugs that prolong QT interval as an unwanted effect: assessing their likelihood of inducing hazardous cardiac dysrhythmias, *Expert Opin. Pharmacother.*, 1, 947–973, 2000.
80. Martin, R.L., McDermott, J.S., Salmen, H.J., Palmatier, J., Cox, B.F., and Gintant, G.A., The utility of hERG and repolarization assays in evaluating delayed cardiac repolarization: influence of multi-channel block, *J. Cardiovasc. Pharmacol.*, 43, 369–379, 2004.
81. McDonald, T.V., Yu, Z., Ming, Z., Palma, E., Meyers, M.B., Wang, K.W., Goldstein, S.A., and Fishman, G.I., A minK-HERG complex regulates the cardiac potassium current I(Kr), *Nature*, 388, 289–292, 1997.
82. Sanguinetti, M.C., Curran, M.E., Zou, A., Shen, J., Spector, P.S., Atkinson, D.L., and Keating, M.T., Coassembly of K(V)LQT1 and minK (IsK) proteins to form cardiac I(Ks) potassium channel, *Nature*, 384, 80–83, 1996.
83. Wakefield, I.D., Pollard, C., Redfern, W.S., Hammond, T.G., and Valentin, J.P., The application of in vitro methods to safety pharmacology, *Fundam. Clin. Pharmacol*, 16, 209–218, 2002.
84. Krasner, L.J., Wendling, W.W., Cooper, S.C., Chen, D., Hellmann, S.K., Eldridge, C.J., McClurken, J.B., Jeevanandam, V., and Carlsson, C., Direct effects of triiodothyronine on human internal mammary artery and saphenous veins, *J. Cardiothorac. Vasc. Anesth.*, 11, 463–466, 1997.
85. Lefer, D.J., Nakanishi, K., Vinten-Johansen, J., Ma, X.L., and Lefer, A.M., Cardiac venous endothelial dysfunction after myocardial ischemia and reperfusion in dogs, *Am. J. Physiol.*, 263, H850–H856, 1992.
86. Sugiyama, A., Kobayashi, M., Tsujimoto, G., Motomura, S., and Hashimoto, K., The first demonstration of CGRP-immunoreactive fibers in canine hearts: coronary vasodilator, inotropic and chronotropic effects of CGRP in canine isolated, blood-perfused heart preparations, *Jpn. J. Pharmacol.*, 50, 421–427, 1989.
87. Voss, H.P., Shukrula, S., Wu, T.S., Donnell, D., and Bast, A., A functional beta-2 adrenoceptor-mediated chronotropic response in isolated guinea pig heart tissue: selectivity of the potent beta-2 adrenoceptor agonist TA 2005, *J. Pharmacol. Exp. Ther.*, 271, 386–389, 1994.
88. Abi-Gerges, N., Small, B.G., Lawrence, C.L., Hammond, T.G., Valentin, J.P., and Pollard, C.E., Evidence for gender differences in electrophysiological properties of canine Purkinje fibres, *Br. J. Pharmacol.*, 142, 1255–1264, 2004.
89. Gintant, G.A., Limberis, J.T., McDermott, J.S., Wegner, C.D., and Cox, B.F., The canine Purkinje fiber: an in vitro model system for acquired long QT syndrome and drug-induced arrhythmogenesis, *J. Cardiovasc. Pharmacol.*, 37, 607–618, 2001.
90. Lu, H.R., Vlaminckx, E., Teisman, A., and Gallacher, D.J., Choice of cardiac tissue plays an important role in the evaluation of drug-induced prolongation of the QT interval in vitro in rabbit, *J. Pharmacol. Toxicol. Methods*, 52, 90–105, 2005.

91. Antzelevitch, C., Shimizu, W., Yan, G.X., and Sicouri, S., Cellular basis for QT dispersion, *J. Electrocardiol.*, 30 Suppl., 168–175, 1998.
92. Hondeghem, L.M., Carlsson, L., and Duker, G., Instability and triangulation of the action potential predict serious proarrhythmia, but action potential duration prolongation is antiarrhythmic, *Circulation*, 103, 2004–2013, 2001.
93. Eckardt, L., Haverkamp, W., Mertens, H., Johna, R., Clague, J.R., Borggrefe, M., and Breithardt, G., Drug-related torsades de pointes in the isolated rabbit heart: comparison of clofilium, d,l-sotalol, and erythromycin, *J. Cardiovasc. Pharmacol.*, 32, 425–434, 1998.
94. Eckardt, L., Haverkamp, W., Borggrefe, M., and Breithardt, G., Experimental models of torsade de pointes, *Cardiovasc. Res.*, 39, 178–193, 1998.
95. Miyoshi, K., Taniguchi, M., Seki, S., and Mochizuki, S., Effects of magnesium and its mechanism on the incidence of reperfusion arrhythmias following severe ischemia in isolated rat hearts, *Cardiovasc. Drugs Ther.*, 14, 625–633, 2000.
96. Hill, A.J., Laske, T.G., Coles, J.A. Jr, Sigg, D.C., Skadsberg, N.D., Vincent, S.A., Soule, C.L., Gallagher, W.J., and Iaizzo, P.A., In vitro studies of human hearts, *Ann. Thorac. Surg.*, 79, 168–177, 2005.
97. Takahara, A., Dohmoto, H., Yoshimoto, R., Sugiyama, A., and Hashimoto, K., Utilization of telemetry system to assess the cardiovascular profile of AH-1058, a new cardioselective Ca2+ channel blocker, in conscious dogs, *Jpn. J. Pharmacol.*, 85, 331–334, 2001.
98. Sato, K., Kandori, H., and Sato, S., Evaluation of a new method using telemetry for monitoring the left ventricular pressure in free-moving rats, *J. Pharmacol. Toxicol. Methods*, 31, 191–198, 1994.
99. Nekooeian, A.A. and Tabrizchi, R., Effects of adenosine a2 receptor agonist, cgs 21680, on blood pressure, cardiac index and arterial conductance in anaesthetized rats, *Eur. J. Pharmacol.*, 307, 163–169, 1996.
100. Ozkanlar, Y., Nishijima, Y., da Cunha, D., and Hamlin, R.L., Acute effects of tacrolimus (FK506) on left ventricular mechanics, *Pharmacol. Res.*, 52, 307–312, 2005.
101. Tsusaki, H., Yonamine, H., Tamai, A., Shimomoto, M., Iwao, H., Nagata, R., and Kito, G., Evaluation of cardiac function in primates using real-time three-dimensional echocardiography as applications to safety assessment, *J. Pharmacol. Toxicol. Methods*, 52, 182–187, 2005.
102. Hanton, G., Gautier, M., Bonnet, P., and Herbet, A., Effect of milrinone on echocardiographic parameters after single dose in Beagle dogs and relationship with drug-induced cardiotoxicity, *Toxicol. Letts.*, 155, 307–317, 2005.
103. Wakefield, I.D., March, J.E., Kemp, P.A., Valentin, J.P., Bennett, T., and Gardiner, S.M., Comparative regional haemodynamic effects of the nitric oxide synthase inhibitors, S-methyl-L-thiocitrulline and L-NAME, in conscious rats, *Br. J. Pharmacol.*, 39, 1235–1243, 2003.
104. van der Linde, H., Van de Water, A., Loots, W., Van Deuren, B., Lu, H.R., Van Ammel, K., Peeters, M., and Gallacher, D.J., A new method to calculate the beat-to-beat instability of QT duration in drug-induced long QT in anesthetized dogs, *J. Pharmacol. Toxicol. Methods*, 52, 168–177, 2005.
105. Sugiyama, A., Ishida, Y., Satoh Y. et al., Electrophysiological, anatomical and histological remodeling of the heart to AV block enhances susceptibility to arrhythmogenic effects of QT prolonging drugs, *Jpn. J. Pharmacol.*, 88, 341–350, 2002.
106. Chiba, K., Sugiyama, A., Hagiwara, T., Takahashi, S.I., Takasuna, K., and Hashimoto, K., In vivo experimental approach for the risk assessment of fluoroquinolone antibacterial agents-induced long QT syndrome, *Eur. J. Pharmacol.*, 486, 189–200, 2004.
107. Collister, J.P. and Osborn, J.W., The chronic infusion of hexamethonium and phenylephrine to effectively clamp sympathetic vasomotor tone: a novel approach, *J. Pharmacol. Toxicol. Methods*, 42, 135–147, 1999.
108. AbdelRahman, A.R., Inadequate blockade by hexamethonium of the baroreceptor heart rate response in anesthetized and conscious rats, *Arch. Int. Pharmacodyn. Ther.*, 297, 68–85, 1989.
109. Verwaerde, P., Senard, J.M., Mazerolles, M., Tran, M.A., DamaseMichel, C., Montastruc, J.L., and Montastruc, P., Spectral analysis of blood pressure and heart rate, catecholamine and neuropeptide Y plasma levels in a new model of neurogenic orthostatic hypotension in dog, *Clin. Auton. Res.*, 6, 75–82, 1996.

110. Harada, T., Abe, J., Shiotani, M., Hamada, Y., and Horri, I., Effect of autonomic nervous function on QT interval in dogs, *J. Toxicol. Sci.*, 30, 229–237, 2005.
111. Mangin, L., Swynghedauw, B., Benis, A., Thibault, N., Lerebours, G., and Carre, F., Relationships between heart rate and heart rate variability: study in conscious rats, *J. Cardiovas. Pharmacol.*, 32, 601–607, 1998.
112. Bedford, T.G. and Dormer, K.J., Arterial hemodynamics during head-up tilt in conscious dogs, *J. Appl. Physiol.*, 65, 1556–1562, 1988.
113. Humphrey, S.J. and McCall, R.B., A rat model for predicting orthostatic hypotension during acute and chronic antihypertensive drug therapy, *J. Pharmacol. Methods*, 7, 25–34, 1982.
114. Ytrehus, K., The ischemic heart—experimental models, *Pharmacol. Res.*, 42, 193–203, 2000.
115. Wu, K.C. and Lima, J.A.C., Non invasive imaging of myocardial viability: current techniques and future developments, *Circ. Res.*, 93, 1146–1158, 2003.
116. Shebuski, R.J., Bush, L.R., Gagnon, A., Chi, L., and Leadley, R.J. Jr, Development and applications of animal models of thrombosis, *Methods Mol. Med.*, 93, 175–219, 2004.
117. Murphy, D.J., Safety pharmacology of the respiratory system: techniques and study design, *Drug Dev. Res.*, 32, 237–246, 1994.
118. Murphy, D.J., Joran, M.E., and Grando, J.C., A non-invasive method for distinguishing central from peripheral nervous system effects of respiratory depressant drugs in conscious rats, *General Pharmacol.*, 26, 569–575, 1995.
119. Jackson, D.M., Pollard, C.E., and Roberts, S.M., The effect of nedocromil sodium on the isolated rabbit vagus nerve, *Eur. J. Pharmacol.*, 221, 175–177, 1992.
120. Widdicombe, J.G., Clinical significance of reflexes from the respiratory system, *Anesthesiology*, 23, 434–444, 1962.
121. Diamond, L. and O'Donnell, M., Pulmonary mechanics in normal rats, *J. Appl. Physiol.: Resp., Environ., Exercise Physiol.*, 43, 942–948, 1977.
122. Mauderly, J.L., Effect of inhaled toxicants on pulmonary function, in *Concepts in Inhalation Toxicology*, McClellan, R.O. and Henderson, R.F., Eds., Hemisphere Publishing Corp., New York, 1989, pp. 347–401.
123. O'Neil, J.J. and Raub, J.A., Pulmonary function testing in small laboratory animals, *Environ. Health Perspect.*, 56, 11–22, 1984.
124. Mauderly, J.L., The influence of sex and age on the pulmonary function of the beagle dog, *J. Gerontol.*, 29, 282–289, 1974.
125. Douglas, J.S., Dennis, M.W., Ridgeway, P., and Bouhuys, A., Airway dilation and constriction in spontaneously breathing guinea-pigs, *J. Pharmacol. Exp. Therapeutics*, 180, 98–109, 1971.
126. Kenakin, T.P., The classification of drugs and drug receptors in isolated tissues, *Pharmacol. Rev.*, 36, 165–222, 1984.
127. Chilvers, M.A. and O'Callaghan, C., Analysis of ciliary beat pattern and beat frequency using digital high speed imaging: comparison with the photomultiplier and photodiode methods, *Thorax*, 55, 314–317, 2000.
128. Khawaja, A.M., Liu, Y.C., and Rogers, D.F., Effect of fenspiride, a non-steroidal anti-inflammatory agent, on neurogenic mucus secretion in ferret trachea in vitro, *Pulm. Pharmacol. Ther.*, 12, 363–368, 1999.
129. Dietl, P., Haller, T., Mair, N., and Frick, M., Mechanisms of surfactantexocytosis in alveolar type II cells in vitro and in vivo, *News Physiol. Sci.*, 16, 239–243, 2001.
130. Murphy, D.J., Assessment of respiratory function in safety pharmacology, *Fund. Clin. Pharmacol.*, 16, 183–196, 2002.
131. Cimbura, G., Lucas, D.M., Bennett, R.C., Warren, R.A., and Simpson, H.M., Incidence and toxicological aspects of drugs detected in 484 fatally injured drivers and pedestrians in Ontario, *J. Forensenic Sci.*, 27, 855–867, 1982.
132. Weiler, J.M., Bloomfield, J.R., Woodworth, G.G., Grant, A.R., Layton, T.A., Brown, T.L., McKenzie, D.R., Baker, T.W., and Watson, G.S., Effects of fexofenadine, diphenhydramine, and alcohol on driving performance. A randomised, placebo-controlled trial in the Iowa driving simulator, *Ann. Intern. Med.*, 132, 354–363, 2000.
133. Mattsson, J.L., Spencer, P.J., and Albee, R.R., A performance standard for clinical and functional observation battery examination of rats, *J. Am. Coll. Toxicol.*, 15, 239, 1996.

134. Irwin, S., Comprehensive observational assessment: 1a. A systematic, quantitative procedure for assessing the behavioural and physiologic state of the mouse, *Psychopharmacol. (Berl.)*, 13, 222–257, 1968.
135. Haggerty, G.C., Strategies for and experience with neurotoxicity testing of new pharmaceuticals, *J. Am. Coll. Toxicol.*, 10, 677–687, 1991.
136. Trabace, L., Cassano, T., Steardo, L., Pietra, C., Villetti, G., and Kendrick, K.M., Biochemical and neurobehavioural profile of CHF2819, a novel, orally active acetylcholinesterase inhibitor for Alzheimer's disease, *J. Pharmacol. Exp. Therap.*, 294, 187–194, 2000.
137. Redfern, W.S., Strang, I., Storey, S., Heys, C., Barnard, C., Lawton, K., Hammond, T.G., and Valentin, J.P., Spectrum of effects detected in the rate functional observational battery following oral administration of non-CNS targeted compounds, *J. Pharmacol. Toxicol. Methods*, 52, 77–82, 2005.
138. Porsolt, R.D., Lemaire, M., Durmuller, N., and Roux, S., New perspectives in CNS safety pharmacology, *Fundam. Clin. Pharmacol.*, 16, 197–207, 2002.
139. Eddy, N.B. and Leimbach, D., Synthetic analgesics: II-Dithienylbutenyl and dithienylbutylamines, *J. Pharmacol. Exp. Ther.*, 107, 385–393, 1953.
140. Krall, E., Antiepileptic drug development: anticonvulsant drug screening, *Epilepsia*, 19, 409–428, 1978.
141. Reiter, L.R. and McPhail, R.C., Motor activity: a survey of methods with potential use in toxicity testing, *Neurobehav. Toxicol.*, 1, 53–66, 1979.
142. Murphy, D.J., Renninger, J.P., and Gossett, K.A., A novel method for chronic measurement of pleural pressure in conscious rats, *J. Pharmacol. Toxicol. Meth.*, 39, 137–141, 1998.
143. Dunham, N.W. and Miya, T.S., A note on a simple apparatus for detecting neurological deficitin mice and rats, *J. Am. Pharm. Assoc.*, 46, 208–209, 1957.
144. Abou-Donia, M.B., Dechkovskaia, A.M., Goldstein, L.B., Shah, D.U., Bullman, S.L., and Khan, W.A., Uranyl acetate-induced sensorimotor deficit and increased nitric oxide generation in the central nervous system in rats, *Pharmacol. Biochem. Behav.*, 72, 881–890, 2002.
145. Jolkkonen, J., Puurunen, K., Rantakomi, S., Harkonen, A., Haapalinna, A., and Sivenius, J., Behavioural effects of the alpha(2)-adrenoceptor antagonist, atipamezole, after focal cerebal ischemia in rats, *Eur. J. Pharamcol.*, 400, 211–219, 2000.
146. Herr, D.W., Graff, J.E., Derr-Yellin, E.C., Crofton, K.M., and Kodavanti, P.R., Flash-, somatosensory-, and peripheral nerve-evoked potentials in rats perinatally exposed to Aroclor 1254, *Neurotoxicol. Teratol.*, 23, 591–601, 2001.
147. Prusky, G.T., Alam, N.M., Beekman, S., and Douglas, R.M., Rapid quantification of adult and developing mouse spatial vision using a virtual optomotor system, *Investig. Ophthalmol. Visual Sci.*, 45, 4611–4616, 2004.
148. Redfern, W.S. and Williams, A., A re-evaluation of the role of α_2-adrenoceptors in the anxiogenic effects of yohimbine, using the selective antagonist delequamine in the rat, *Br. J. Pharmacol.*, 116, 2081–2089, 1995.
149. Jouvet, M., Biogenic amines and the states of sleep, *Science*, 163, 32–41, 1969.
150. Bammer, C., Pharmacological investigations of neurotransmitter involvement in passive avoidance responding: a review and some new results, *Neurosci. Biobehav. Rev.*, 6, 247–296, 1982.
151. Glick, S.D. and Zimmerberg, D., Amnesic effects of scopolamine, *Behav. Biol.*, 7, 2445–2454, 1972.
152. Morris, R.G.M., Spatial localization does not require the presence of local cues, *Learn. Motiv.*, 12, 239–260, 1981.
153. Olton, D.S., The radial maze as a tool in behavioural pharmacology, *Physiol. Behav.*, 40, 793–797, 1986.
154. Dunnett, S.B., Evenden, J.L., and Iversen, S.D., Delay-dependent short term memory deficits in aged rats, *Psychopharmacology*, 96, 174–180, 1982.
155. Itil, T.M., The discovery of psychotropic drugs by computer-analyzed cerebral bioelectrical potentials (CEEG), *Drug Dev. Res.*, 1, 373–407, 1981.
156. Meert, T.F., Effects of various serontonergic agents on alcohol intake and alcohol preference in Wistar rats selected at two different levels of alcohol preference, *Alcohol Alcoholism*, 28, 157–170, 1993.

157. Goudie, A.J., Harrison, A.A., and Leathley, M.J., Evidence for a dissociation between benzodiazepine withdrawl signs, *Neuro-report*, 4, 295–299, 1993.
158. Brady, J.V. and Fischman, M.W., Assessment of drugs for dependence potential and abuse liability: an overview, in *Behavioural Pharmacology: The Current Status*, Seiden, L.S. and Balster, R.L., Eds., Alan R. Liss, New York, U.S.A., 1985, pp. 361–382.
159. Lal, H., Ed., *Discriminative Stimulus Properties of Drugs*, Plenum Press, New York, U.S.A., 1977.
160. Schechter, M.D. and Calcagnetti, D.J., Trends in place preference conditioning with cross-indexed bibliography 1957–91, *Neurosci. Biobehav. Rev.*, 17, 21–41, 1993.
161. Fountain, S.B. and Teyler, T.J., Suppression of hippocampal slice excitability by 2-, 3- and 4-methylpyridine, *Ecotoxicol. Environ. Safety*, 48, 301–305, 1995.
162. Fliri, A.F., Loging W.T., Thadeio, P.F., and Volkmann, R.A., Biological spectra analysis: linking biological activity profiles to molecular structure, *PNAS*, 102, 261–266, 2005.
163. Henry, D.A., Ostapowicz, G., and Robertson, J., Drugs as a cause of gastrointestinal disease, *Baillieres Clin. Gastroenterol.*, 8, 271–300, 1994.
164. Makins, R. and Ballinger, A., Gastrointestinal side effects of drugs, *Expert Opin. Drug Safety*, 2, 421–429, 2003.
165. Mayer, E.A. and Bradesi, S., Alosetron and irritable bowel syndrome, *Expert Opin. Pharmacother.*, 4, 2089–2098, 2003.
166. Chassany, O., Michaux, A., and Bergmann, J.F., Drug-induced diarrhoea, *Drug Safety*, 22, 53–72, 2000.
167. Lewis, J.H., Gastrointestinal injury due to medicinal agents, *Am. J. Gastroenterol.*, 81, 819–834, 1986.
168. Ghahremani, G.G., Gastrointestinal complications of drug therapy, *Abdom. Imaging*, 24, 1–2, 1999.
169. Gatenby, R.A., The radiology of drug-induced disorders in the gastrointestinal tract, *Sem. Roentgenol*, 30, 62–76, 1995.
170. Kimura, T. and Higaki, K., Gastrointestinal transit and drug absorption, *Biol. Pharm. Bull.*, 25, 149–164, 2002.
171. Gershon, M.D., The enteric nervous system: a second brain, *Hosp. Pract*. (Office Edition), 34, 31–32, 1999.
172. Harrison, A.P., Erlwanger, K.H., Elbrond, V.S., Andersen, N.K., and Unmack, M.A., Gastrointestinal-tract models and techniques for use in safety pharmacology, *J. Pharmacol. Toxicol. Methods*, 49, 187–199, 2004.
173. Andrews, P.L. and Hawthorn, J., The neurophysiology of vomiting, *Baillieres Clin. Gastroenterol.*, 2, 141–168, 1998.
174. Lang, I.M., Noxious stimulation of emesis, *Dig. Dis. Sci.*, 44, 58–63, 1999.
175. Hornby, P.J., Central neurocircuitry associated with emesis, *Am. J. Med.*, 111, 106–112, 2001.
176. Rudd, J.A. and Naylor, R.J., The actions of ondansetron and dexamethasone to antagonise cisplatin-induced emesis in the ferret, *Eur. J. Pharmacol.*, 322, 79–82, 1997.
177. Yamashita, M., Yamashita, M., Tanaka, J., Chagi, K., Takeda, S., Kurihara, T., Takeda, Y., and Fujii, Y., Vomiting induction by ipecac syrup in dogs and ferrets, *J. Toxicol. Sci.*, 22, 409–412, 1997.
178. Fillastre J.P. Détection de la néphrotoxicité médicamenteuse, *J. Pharmacol. (Paris)*, 17(Suppl. I), 41–50, 1986.
179. Thatte, L. and Vaamonde, C.A., Drug-induced nephrotoxicity, *Postgrad. Med.*, 100, 83–100, 1996.
180. Thadani, R., Pascual, M., and Boventre, J.V., Acute renal failure, *N. Engl. J. Med.*, 334, 1448–1460, 1996.
181. Lieberthal, W. and Nigam, S.K., Acute renal failure II. Experiemtnal models of acute renal failure: imperfect but indispensable, *Am. J. Physiol.*, 278, F1–F12, 2000.
182. Zbinden, G., Ed., *Pharmacological Methods in Toxicology*, Pergamon, Elmsford, New York, 1997, pp. 3–6.
183. Luft, J. and Bode, G., Integration of safety pharmacology endpoints into toxicology studies, *Fundam. Clin. Pharmacol.*, 16, 91–103, 2002.
184. Haefely, W., Biological basis of drug-induced tolerance, rebound, and dependence. Contribution of recent research on benzodiazepines, *Pharmacopsychiatry*, 19, 353–361, 1986.

185. McMahon, N., Robinson, N.A., Martel, E., and Valentin, J.P., A method for the long term monitoring of cardiovascular and respiratory function in the Wistar rat, *Toxicology,* 202, 33–127, 2004.
186. Hunter, D., Schofield, J., Gracie, K., Moors, J., Philp, K., Carter, P., Prior, H., Valentin, J.P., and Hammond, T.G., Use of a non-invasive telemetry system (EMKA) for functional cardiovascular endpoints in toxicology studies, *J. Pharmacol. Toxicol. Methods*, 49, 234, 2004.
187. Tattersall, M.L., Dymond, M., Hammond, T., and Valentin, J.P., Correction of QT values to allow for increases in heart rate in conscious beagle dogs in toxicology assessment, *J. Pharmacol. Toxicol. Method*, 53, 11–19, 2006.
188. Detweiller, D.K., Electrocadiography in toxicological studies, in *Comprehensive Toxicology*, Snipes, I.G., McQueen, C.A., and Gandolfi, A.J., Eds., Pergamon Press, New York, U.S.A., 1997.
189. Opie, L.H., *The Heart Physiology, From Cell to Circulation*, Lippincott-Raven, Philadelphia, 1998.
190. Hamlin, R.L., Non-drug-related electrocardiographic features in animal models in safety pharmacology, *J. Pharmacol. Toxicol. Methods*, 52, 60–76, 2005.
191. LeBel, C.P. and Foss, J.A., Use of a rodent neurotoxicity screening battery in the preclinical safety assessment of recombinant-methionyl human brain-derived neurotrophic factor, *Neurotoxicology*, 17, 851–863, 1996.
192. Horner, S.A., Gould, S., Noakes, J.P., Rattray, N.J., Allen, S.L., Zotova, E., and Arezzo, J.C., Lack of neurotoxicity of the vascular targeting agent ZD6126 following repeated i.v. dosing in the rate, *Mol. Cancer Ther.*, 3, 783–791, 2004.
193. Chapillon, P., Lalonde, R., Jones, N., and Caston, J., Early development of synchronised walking on the rodarod in rats — Effects of training and handling, *Behav. Brain Res.*, 93, 77–81, 1998.
194. Holliday, T.A., Nelson, H.J., Williams, D.C., and Willits, N., Unilateral and bilateral brainstem auditory-evoked response abnormalities in 900 Dalmatian dogs, *J. Vet. Intern. Med.*, 6, 166–174, 1992.
195. Colley, J.C., Edwards, J.A., Heywood, R., and Purser, D., Toxicity studies with quinine hydrochloride, *Toxicology*, 54, 219–226, 1989.
196. Redfern, W.S., 2005, unpublished observations.
197. Moser, V.C., Becking, G.C., MacPhail, R.C., and Kulig, B.M., The IPCS collaborative study on neurobehavioral screening methods, *Fundam. Appl. Toxicol.*, 35, 143–151, 1997.
198. Ando, K., Ikeda, H., Yamamoto, K., and Sagami, F., Dose (concentration)-response analysis of drug-induced QT interval prolongation in conscious monkeys (JPMA QT prodact), *J. Pharmacol. Toxicol. Methods*, 49, 220, 2004.
199. Ando, K., Sugiyama, A., Satoh, Y., Nakamura, Y., and Hashimoto, K., Predicting drug-induced QT prolongation using a new in vivo animal model: comparison of risperidone and olanzapine, *J. Pharmacol. Toxicol. Methods*, 49, 221, 2004.
200. Miyazaki, H., Kitayama, T., Sekiya, K., Haruna, M., Mino, T., Suganami, H., Watanabe, H., and Yamamoto, K., Individual QT-RR correction and sensitivity to detect drug-induced changes in QT interval for canine telemetry assay, *J. Pharmacol. Toxicol. Methods*, 49, 224, 2004.
201. Miyazaki, H., Kitayama, T., Tashibu, H., Ando, K., Yamamoto, K., and Sagami, F., Comparison of the sensitivity between canine telemetry assay and isofluorane-anaesthetised model (JPMA QT Prodact), *J. Pharmacol. Toxicol. Methods*, 49, 224, 2004.
202. De Bruin, M.L., Pettersson, M., Meyboom, R.H.B., Hoes, A.W., and Leujkens, H.G.M., Anti-hERG activity and the risk of drug-induced arrhythmias and sudden death, *Eur. Heart J.*, 26, 590–597, 2005.
203. Kinter, L.B. and Johnson, D.K., Remote monitoring of experimental endpoints in animals using radiotelemetry and bioimpedance technologies, in *Proceedings of the International Conference on Humane Endpoints in Animal Experiments for Biomedical Research*, Hendriksen, C.F.M. and Morton, D.B., Eds., The Royal Society of Medicine Press, London, 1999, pp. 58–65.
204. Kramer, K. and Kinter, L.B., Evaluation and application of radiotelemetry in small laboratory animals, *Physiol. Genomics*, 13, 197–205, 2003.
205. Kramer, K., Mills, P.A., Kinter, L.B., and Brockway, B.P., History of laboratory telemetry and the state of the art of fully implanted radiotelemetry for monitoring laboratory animals, *Lab. Anim.*, 27, 40–46, 1998.

206. Davila, J.C., Cezar, G.G., Thiede, M., Strom, S., Miki, T., and Trosko, J., Use and application of stem cells in toxicology, *Toxicol. Sci.*, 79, 214–223, 2004.
207. Rubinstein, A.L., Zebrafish: from disease modeling to drug discovery, *Curr. Opin. Drug Discov. Dev.*, 6, 218–223, 2003.
208. Zon, L.I. and Peterson, R.T., In vivo drug discovery in the zebrafish, *Nat. Rev. Drug Discov.*, 4, 35–44, 2005.
209. Parng, C., In vivo zebrafish assays for toxicity testing, *Curr. Opin. Drug Discov. Dev.*, 8, 100–106, 2005.
210. Spitsbergen, J.M. and Kent, M.L., The state of the art of the zebrafish model for toxicology and toxicologic pathology research — Advantages and current limitations, *Toxicol. Pathol.*, 31, 62–87, 2003.
211. Farber, S.A., Pack, M., Ho, S.Y., Johnson, I.D., Wagner, D.S., Dosch, R., Mullins, M.C., Hendrickson, H.S., Hendrickson, E.K., and Halpern, M.E., Genetic analysis of digestive physiology using fluorescent phospholipid reporters, *Science*, 292, 1385–1388, 2001.
212. Wallace, K.N. and Pack, M., Unique and conserved aspects of gut development in zebrafish, *Dev. Biol.*, 255, 12–29, 2003.
213. Baraban, S.C., Taylor, M.R., Castro, P.A., and Baier, H., Pentylenetetrazole induced changes in zebrafish behavior, neural activity and c-fos expression, *Neuroscience*, 131, 759–768, 2005.
214. Pasini, A. and Wilkinson, P.G., Stabilizing the regionalisation of the developing vertebrate central nervous system, *Bioassays*, 24, 427–438, 2002.
215. Gerlai, R., Lahav, M., Guo, S., and Rosenthal, A., Drinks like a fish: Zebra fish (*Danio rerio*) as a behaviour genetic model to study alcohol effects, *Pharmacol. Biochem. Behav.*, 67, 773–792, 2000.
216. Gerlai, R., Zebra fish: an unchartered behaviour genetic model, *Behav. Genet.*, 33, 461–468, 2003.
217. Zhdanova, I.V., Wang, S.Y., Leclair, O.U., and Danilova, N.P., Melatoinin promotes sleep-like state in zebrafish, *Brain Res.*, 903, 263–268, 2001.
218. Kimmel, C. et al., The development and behavioral characteristics of the startle response in the zebra fish, *Dev. Psychobiol.*, 7, 47–60, 1974.
219. Bang, P.I., Yelick, P.C., Malicki, J.J., and Sewell, W.F., High-throughput behavioral screening method for detecting auditory response defects in zebrafish, *J. Neurosci. Methods*, 118, 177–187, 2002.
220. Zeddies, D.G. and Fay, R.R., Development of the acoustically evoked behavioral response in zebrafish to pure tones, *J. Exp. Biol.*, 208, 1363–1372, 2005.
221. Whitfield, T.T., Zebrafish as a model for hearing and deafness, *Neurobiology*, 53, 157–171, 2002.
222. Fritsche, R., Schwerte, T., and Pelster, B., Nitric oxide and vascular reactivity in developing zebrafish, *Danio rerio*, *Am. J. Physiol. Regul. Integr. Comp. Physiol.*, 279, R2200–R2207, 2000.
223. Hu, N., Sedmera, D., Yost, H.J., and Clark, E.B., Structure and function of the developing zebrafish heart, *Anat. Rec.*, 250, 148–157, 2000.
224. Milan, D.J., Peterson, T.A., Ruskin, J.N., Peterson, R.T., and MacRae, C.A., Drugs that induce repolarization abnormalities cause bradycardia in zebrafish, *Circulation*, 107, 1355–1358, 2003.
225. Ulrike, L., Gabriele, V., and Thomas, W., Zebrafish embryos express an orthologue of HERG and are sensitive toward a rank of QT-prolonging drugs inducing severe arrhythmia, *Toxicol. Appl. Phamacol.*, 193, 370–382, 2003.
226. Schwerte, T., Uberbacher, D., and Pelster, B., Non-invasive imaging of blood cell concentration and blood distribution in zebrafish *Danio rerio* incubated in hypoxic condition in vivo, *J. Exp. Biol.*, 206, 1299–1307, 2003.
227. Grosser, T., Yusuff, S., Cheskis, E., Pack, M.A., and FitzGerald, G.A., Developmental expression of functional cyclooxygenases in zebrafish, *Proc. Natl. Acad. Sci.*, 99, 8418–8423, 2002.
228. Langenua, D.M., Ferrando, A.A., Traver, D., Kutok, J.L., Henzel, J.P., Kanki, J.P., Zon, L.I., Look, A.T., and Trede, N.S., In vivo tracking of T cell development ablation and engraftment in transgenic zebrafish, *Proc. Natl. Acad. Sci. USA*, 101, 7369–7374, 2004.
229. Hsieh, D.J. and Liao, C.F., Zebrafish M2 muscarinic acetylcholine receptor: cloning, pharmacological characterization, expression patterns and roles in embryonic bradycardia, *Br. J. Pharmacol.*, 137, 782–792, 2002.
230. Egan, J.E., Zlokarnik, G., and Grootenhuis, P.D.J., In silico prediction of drug safety: despite progress there is abundant room for improvement, *Drug Discov. Today Technol.*, 1, 381–387, 2004.

231. Ekins, S., Crumb, W.J., Sarazan, R.D., Wikel, J.H., and Wrighton, S.A., Three-dimensional quantitative structure activity relationship of human ether-a-go-go-related gene potassium channel, *J. Pharmacol. Exp. Ther.*, 301, 427–434, 2002.
232. Cavalli, A., Poluzzi, E., De Ponti, F., and Recanatini, M., Toward a pharmacophore for drugs inducing the long QT syndrome: insights from a CoMFA study of HERG K+ channel blockers, *J. Med. Chem.*, 45, 3844–3853, 2002.
233. Pearlstein, R.A., Vaz, R.J., Kang, J., Chen, X.L., Preobrazhenskaya, M., Shchekotikhin, A.E., Korolev, A.M., Lysenkova, L.N., Miroshnikova, O.V., Hendrix, J., and Rampe, D., Characterisation of HERG potassium channel inhibition using CoMSiA 3D QSAR and homology modelling approaches, *Bioorg. Med. Chem. Lett.*, 13, 1829–1835, 2003.
234. Zolotoy, A.B., Plouvier, B.P., Beatch, G.B., Hayes, E.S., Wall, R.A., and Walker, M.J., Physicochemical determinants for drug induced blockade of HERG potassium channels: effects of charge and charge shielding, *Curr. Med. Cardiovasc. Hematol. Agents*, 1, 225–241, 2003.
235. Netzer, R., Bischoff, U., and Ebneth, A., HTS techniques to investigate the potential effects of compounds on cardiac ion channels at early-stages of drug discovery, *Curr. Opin. Drug Discov. Dev.*, 6, 462–469, 2003.
236. Garny, A., Noble, D., and Kohl, P., Dimensionality in cardiac modelling, *Prog. Biophys. Mol. Biol*., 87, 47–66, 2005.
237. Noble, D., Modelling the heart, *Physiology*, 19, 191–197, 2004.
238. Ten Tusscher, K.H., Noble, D., Noble, P.J., and Panfilov, A.V., A model for human ventricular tissue, *Am. J. Physiol. Heart Circ. Physiol*., 286, H1573–H1589, 2004.
239. Noble, D., Modelling the heart — From genes to cells to the whole organ, *Science*, 295, 1678–1682, 2002.
240. Norinder, U., Sjoberg, P., and Osterberg, T., Theoretical calculation and prediction of brain-blood partitioning of organic solutes using molsurf parametrization and PLS statistics, *J. Pharm. Sci.*, 87, 952–959, 1998.
241. Sarapa, N., Early human microdosing to reduce attrition in clinical drug development, *Am. Pharm. Outsourcing*, 4, 42–46, 2003.
242. Dixit, R., What non-clinical toxicology and safety pharmacology data are needed to accelerate phase I–II clinical trials?, *Am. Pharm. Outsourcing*, 5, 30–37, 2004.
243. Anon., Innovation or stagnation, challenge and opportunity on the critical path to new medicinal products, March 2004.
244. Calamari, D., Strategic survey of therapeutic drugs in the rivers Po and Lambo in Northern Italy, *Environ. Sci. Technol.*, 37, 1241–1248, 2003.
245. Huggett, D.B., Khan, I.A., Foran, C.M., and Schlenk, D., Determination of beta-adrenergic receptor blocking pharmaceuticals in United States wastewater effluent, *Environ. Pollut.*, 121, 199–205, 2003.
246. Kopin, D.W., Furlong, E.T., Meyer, M.T., Thurman, E.M., Zaugg, S.D., Barber, L.B., and Buxton, H.T., Pharmaceuticals, hormones, and other organic wastewater contaminants in U.S. streams 1999–2000: a national reconnaissance, *Environ. Sci. Technol.*, 36, 1202–1211, 2002.
247. Anon., In vivo pharmacology training group. The fall and rise of in vivo pharmacology, *Trends Pharmacol. Sci.*, 23, 13–18, 2002.
248. Jobe, P.C., Adams-Curtis, L.E., Burks, T.F., Fuller, R.W., Peck, C.C., Ruffolo, R.R., Snead III, O.C., and Woosley, R.L., The essential role of the integrative biomedical sciences in protecting and contributing to the health and well being of our nation, *Pharmacologist*, 40, 32–37, 1998.

14

Nonclinical Drug Safety Assessment

Frederick E. Reno

CONTENTS

14.1 Introduction

As most individuals involved in drug development know only too well, the ultimate safety assessment of any new drug comes from extensive clinical studies. However, prior to its first introduction into humans, and generally continuing through the clinical

program, an extensive series of nonclinical studies, using laboratory animals as well as an emerging array of *in vitro* studies, were carried out to support the safe use of the proposed pharmaceutical products in clinical trials as well as to support the new drug application (NDA) concerning certain label requirements. The discovery stage has shown, using *in vivo* or *in vitro* studies, the pharmacological profile of the new drug and a rationale for the potential efficacy of the drug. Different companies use different strategies with regard to the types and sufficiency of data (potency, potential therapeutic margin, etc.) and information (marketplace, competition, etc.) needed to declare a new chemical entity (NCE) a development candidate. At this stage, the candidate enters development, a multidisciplinary activity encompassing the fields of toxicology, drug metabolism, formulation development (for both nonclinical and clinical studies), safety pharmacology, and multiple other disciplines. Many of these disciplines will be discussed in other chapters. The knowledge of submolecular and genetically controlled mechanisms is growing daily, and the role of genomics in safety evaluation is expanding.[1,2] This chapter will focus on nonclinical toxicology and its role in this very complex process, focusing on the process in the United States and finally discussing the more global perspective through the International Conference on Harmonization (ICH) and pointing out some jurisdictional differences.

14.2 History of Nonclinical Guidelines

14.2.1 Food and Drug Administration

The Federal Food, Drug and Cosmetic Act of 1906 was the first major national legislation for the United States dealing with pharmaceutical products and attempted to place some control over the interstate distribution of drugs. Interestingly, there were no provisions that addressed the more important aspects that we know today, that is, proof of safety and efficacy. And certainly there were no provisions related to nonclinical safety assessment. In 1956, Dr Arnold Lehman, a half century after the passage of the act, offered some general approaches at various meetings related to the conduct of preclinical studies that might establish the basis for the safe use of new drugs being evaluated in man. These suggestions encompassed several pages of outlines of some studies for consideration. It was not until the potential tragedy of thalidomide was averted in the United States that the Congress of the country passed the Kefauver–Harris amendments to the Act in 1962 that, for the first time, required, among other things, proof of safety and efficacy. A small cadre of individuals within the Food and Drug Administration (FDA), as it was known by that time, worked together, and the outcome was that preclinical requirements came into existence. In the ensuing years, representatives of the pharmaceutical industry were requesting that the agency establish more firm guidelines, which would support the Investigational New Drug Application (IND) now required by the amendments to the Act. In 1966, Dr Edwin Goldenthal of the FDA wrote what became known, even to this day, as the "Goldenthal Guidelines"[3] and was published in an FDA periodical called "FDA Papers." Therein, studies were outlined in somewhat more detail, and included general guidelines for animal toxicity studies that related the duration of animal studies to the duration of clinical treatment from IND to the NDA, which describe the sequence of reproduction and teratology studies with the categorization of Segments I, II, and III, and noted the requirement for 7-year studies in

dogs and 10-year studies in monkeys for oral contraceptives (based upon concerns regarding breast cancer). This latter requirement has since been rescinded. Dr William D'Aguanno of FDA was also instrumental in delineating study designs for nonclinical studies.

The 1970s brought the era of major FDA inspections of both pharmaceutical companies and contract research organizations that uncovered some inappropriate practices in the conduct of toxicology studies. Following congressional hearings and much interaction with representatives of the pharmaceutical industry, the FDA promulgated Good Laboratory Practice (GLP) regulations that became final in June 1979. From that point on, all nonclinical studies that were conducted to support a clinical research or marketing application had to be carried out in compliance with these regulations and the study reports were audited by the performing laboratory's quality assurance unit (QAU). The existence of a QAU was one of the mandated provisions of the GLP. Pharmacology studies and certain preliminary toxicology studies are exempted from these regulations.

14.2.2 International Conference on Harmonization

The International Conference on Harmonization (ICH) was formed from the three major regulatory jurisdictions (United States, Europe, and Japan) with a goal of establishing a common set of guidelines that would address many aspects of the entire drug development process, including nonclinical studies. These meetings, which began in the 1980s, have resulted in the issuance of many guidelines beginning in 1991 to the present time (either in draft or final form) for nonclinical studies and are applicable in all three jurisdictions. These will be referenced throughout this chapter. In addition, each of the jurisdictions has also issued other guidelines that can be considered specific to their region. The U.S. FDA issues "Guidance for Industry" documents and are available on the agency's web site. The European Medicines Agency (EMEA) was formed in 1995 and is a decentralized body of the European Union and coordinates the evaluation of pharmaceutical products throughout the European Union. The EMEA has also issued some guidelines.

14.2.3 Global Development Considerations

It has occasionally been the practice for American companies to initiate initial phase 1 studies abroad. This is because these studies can often be initiated outside the United States with little regulatory involvement, and with safety considerations being reviewed by an Institutional Review Board or Ethics Committee. Certainly, this is now not always the case, as the European Clinical Trials Directive of 2004 now requires a Clinical Trials Authorization (CTA) and submission to the regulatory authorities in the United Kingdom (Medicines and Healthcare Products Regulatory Agency, or MHRA). Nevertheless, initiation of clinical trials without specific regulatory submissions is still allowed in some countries of Europe. Phase 1 studies can be initiated more rapidly and thus obtain information on preliminary safety and pharmacokinetic data earlier. This approach requires a different approach than would a strategy requiring the filing of an IND or CTA.

No pharmaceutical development program is undertaken without consideration given to eventual global registration and marketing. Although this may be a matter of timing, nonclinical programs must take into account whatever nuances may exist in the various jurisdiction.

14.3 Nonclinical Consistent with Clinical Development

14.3.1 Timing of Clinical Studies

The nature, timing, and extent of the initial nonclinical toxicology effort must be consistent with the clinical development plan that it must support. The ICH guidelines further specify the extent and duration of nonclinical studies that are required to initiate or continue clinical studies. Therefore, it is important that the clinical development plan, at least the initial stages, be clearly delineated so that clinical studies are not delayed owing to the lack of appropriate nonclinical studies.

14.3.2 Enrollment of Women

The inclusion of women in the clinical program, where the indication of the drug is not gender-specific, should take place as early in phase 2 as possible. Obviously, drugs that are designed for use in women only requires enrollment of women in all phases. Since the thalidomide episode, teratology studies and possibly other reproduction studies have been required prior to enrollment of large numbers of women in clinical studies. Early clinical studies, even in phase 1, have enrolled women without teratology studies if the women have been postmenopausal or otherwise reproductively incapable (such as practicing two forms of contraception). However, the timing of the enrollment of women needs to be understood well in advance so that the lack of appropriate nonclinical reports does not hinder clinical development.

14.4 Formulation Considerations

The identification of a formulation to be used in nonclinical studies can depend upon a variety of factors, including route of administration, solubility of the drug, the pH of the drug in solution, the dose that needs to be delivered, etc. This is often not a problem if the marketed formulation is tablets or capsules. In these cases, drugs can be given to animals in solutions, suspensions, and, for some nonrodents, dry filled capsules. It is with parenteral formulations that many problems are encountered, as often the active ingredient is insoluble and the selection of a physiologically suitable vehicle can be challenging. Frequently, particularly with some insoluble anticancer drugs, low concentrations are necessary, necessitating the use of continuous infusion studies to increase the dose per day.

The type of formulation can affect the pharmacokinetics of the drug and thus can alter the toxicological profile, making comparison of animal and human pharmacokinetics occasionally difficult.

Impurities are a critical focus, particularly in early stage studies. Furthermore, scaling up the manufacturing processes can result in active ingredients with different impurities. There are ICH guidelines[4] that address impurities in drug substances and their qualification if they exceed guidance levels. Another important factor is the GLP requirement that every batch of drug used in nonclinical studies must have a certificate of analysis that clearly specifies the purity level as well as the quantities of impurities, which may

include residual solvents, unreacted starting materials, entities formed during manufacture, degradants, etc. The impurities must be reviewed in terms of the potential contribution that they can make to toxic effects, which may be manifested in the nonclinical studies. Furthermore, animals must be exposed to any impurity to which humans are exposed.

Of equal importance is the stability of the drug in the nonclinical formulation. This can determine the frequency at which nonclinical formulations must be prepared (daily, weekly, etc.). Orally administered drugs must be resistant to degradation by gastric acids to the extent possible and must be stable in the formulation itself. Analytical methods must be available in the very early stages of drug development to confirm stability, homogeneity, and potency of dosing formulations.

The amount of active ingredient required to carry out nonclinical studies is frequently substantial compared to that needed for initial clinical studies. While many biologically derived drugs may require relatively small quantities owing to the potency of the material or the limited number of nonclinical studies needed, a typical program needed for IND-enabling studies for drugs that are relatively nontoxic may require 2 to 3 kg of active ingredient. For many companies, this can be difficult either from a manufacturing standpoint (small quantities synthesized prior to scale-up) or cost.

14.5 Analytical Requirements

Prior to the implementation of a nonclinical program, there are two major areas of analytical needs.

14.5.1 Dose Confirmation

Good Laboratory Practice regulations require that all active ingredients mixed with a carrier or vehicle must have some or all (depending on the duration of the study) of the formulations analyzed for potency compared with the nominal concentration. For short-term studies, the analyses may have to be more frequent than longer term studies (e.g., quarterly for carcinogenicity studies). This requires the availability of a validated analytical method at the earliest stages of development. Although not directly specified in the regulations, the FDA is moving to an interpretation that much of this analytical work must also be performed under GLPs. Since analytical work for smaller companies is frequently contracted to analytical laboratories, this has occasionally been a problem, as many analytical labs are not able to fully comply with all aspects of the GLPs.

14.5.2 Assays for Drug in Plasma

Toxicokinetics has become a critically important component of any nonclinical program (see discussion in Section 14.10). Current ICH guidelines require the determination of animal pharmacokinetics at all dose levels administered on at least 2 days (beginning and end) during a nonclinical toxicology study.[5] Similarly, this requires the development of a validated analytical method for the determination of parent drug (and possible major metabolites).

14.6 Species Selection

The major purpose of nonclinical studies is to identify the toxicological properties of the proposed candidate so as to assure a reasonable margin of safety in man on the basis of either exposure (AUC) or dose (mg/kg/day or mg/m^2/day). An essential underlying premise is to use animals that pharmacologically and pharmacokinetically handle the drug in a way similar to man. Convenience and ease of handling have resulted in the majority of these studies being performed in rodents (rats and mice) and nonrodents (dogs and primates). One teratology study is done in rabbits. In the past, following the completion of phase 1 studies, it was determined if our selection of these species was correct, or at least close. With the validation of *in vitro* metabolism studies, it is often possible to select the appropriate species before the IND-enabling studies are even initiated. In these studies, the drug is incubated with hepatic microsomes, hepatocytes or liver slices, and metabolites identified, either qualitatively or quantitatively. When candidate drugs are proteins, animal- or human-derived monoclonal antibodies, or sometimes proteins or peptides with some homology to human proteins, then antibody formation may be a major issue and may dictate the choice of species. For example, experience has shown in some cases that only the chimpanzee does not develop neutralizing antibodies to the drug and the species has been used as the only nonclinical model. Biological products are often evaluated for toxicity in the same model as was used to determine efficacy.

Topical formulations are another special case. Over time, it has been shown that the minipig has a skin structure that is quite similar to humans, and that species is now used commonly as the nonrodent model. These types of formulations also require local irritation studies where guinea pigs are used to determine delayed contact sensitization. Selection of the animal species for the nonclinical program is often not straightforward.

14.7 Supporting the Investigational New Drug Application

Initial clinical trials are usually designed to study tolerability and to provide initial pharmacokinetic assessments. An ICH guideline identifies studies that are needed to support various stages of clinical development.[6] For an IND where the drug will be given as single doses to normal subjects, the IND-enabling studies usually involve 14 to 28 days of repeated administration. Under some circumstances, single-dose animal studies can be used if the clinical studies are limited to single doses and the full spectrum of toxicological parameters are included in the protocol.[7] A 28-day program may allow coverage for early phase 2 studies if proof of principle can be determined within that time period. Thus, the medium range strategy for clinical studies must be determined early on so that the nonclinical program can be designed appropriately. The FDA recently issued a guidance that allows for exploratory clinical studies to be supported by an IND with somewhat less stringent requirements.[8]

The preliminary evaluation of the safety assessment of any new drug requires multiple studies, some of which evaluate general and multiple endpoints (such as toxicity studies) and others evaluate more specific and defined endpoints (such as mutagenicity studies and safety pharmacology studies). Drugs that are derived from a biological origin, such as proteins, monoclonal antibodies, or drugs produced by recombinant technology (or what are generally referred to as "biotechnology products"), present additional problems that

require a significantly modified approach. The ICH guidelines recognize that unique approaches may be needed, and has addressed this in a further guideline.[9] Terrell and Green[10] have pointed out the problems that can be posed for the toxicologist in evaluating biotechnology products. The FDA has published guidelines that outline the requirements and types of studies necessary to support an IND.[11] This latter document focuses more on the extent of study documentation required than the study types, and allows for data to be submitted that are not in final report form. This FDA document covers both xenobiotic as well biotechnology-derived drugs. There are two types of guidelines that must be considered in initiating the nonclinical program. The first relates to the types of studies required and the second to protocol requirements for the studies themselves.

The following sections briefly describe the studies needed to support an IND in the United States and other jurisdictions. Additional specialized studies might be needed in order to study the potential for an effect that might be characteristic of drug in the particular class in questions (e.g., antibody determinations for some biological products, neurotoxicity studies for drugs acting on the central nervous system). Also, some regulatory jurisdictions may have additional requirements.

The U.S. FDA has issued a guidance document that assists both the toxicologist and the clinician to propose the first dose in man.[12] The starting dose becomes a fraction (often 10%) of the lowest no-adverse-effect level (NOAEL) in the rodent or nonrodent repeated dose studies. This is an oversimplistic statement and the reader is referred to the specifics of the guidance.

Anticancer drugs represent another departure from the standard IND approach in the United States and other jurisdictions. While many companies do follow the IND guidelines because of some other possible planned indications, another guideline is followed[13] owing to the nature of these agents and the modified and often cyclical nature of the dosing schedule. The animal studies generally follow the planned cyclic regimen of the clinical program, and mutagenicity and safety pharmacology studies are usually not required if initial clinical studies in the United States are performed in cancer patients.

14.7.1 Acute Toxicity Studies

Single-dose studies in animals is an important first step in establishing a safety profile. While the calculation of an LD_{50} is no longer required, it is important to study a range of doses that will allow the identification of a dose without drug-related effects, a dose that produces some level of exaggerated effect (not necessarily death) that helps identify potential side effects, and other doses in between that help characterize the tolerability of the drug. These studies can be designed using "up-and-down" procedures or even redose the same animals to reduce the number of animals required. These studies should provide sufficient information to select doses for the first repeated dose studies. As mentioned previously, a guideline is available addressing acute studies.[7]

14.7.2 Repeated Dose Studies

Repeated dose studies are designed to identify safe levels of the drug given on a repeated basis and establish a good safety margin compared to the anticipated clinical doses. The route of administration should be the same as that planned in man. The types and duration of specific studies, and which ones are needed relative to different stages of clinical development, were mentioned previously (ICH M3), and the reader is referred to Tables 14.1 and 14.2. Protocols must specify the number of animals per group, numbers of groups and experimental procedures to be carried out, and standard versions of these

TABLE 14.1

Duration of Repeated Dose Toxicity Studies to Support Phase 1 and Phase 2 Clinical Trials in the EU, and Phase 1, 2, and 3 Clinical Trials in the United States and Japan[a]

	Minimum Duration of Repeated Dose Toxicity Studies	
Duration of Clinical Trials	Rodents	Nonrodents
single dose	2 weeks[b]	2 weeks
≤2 weeks	2 weeks[b]	2 weeks
≤1 month	1 month	1 month
≤3 months	3 months	3 months
≤6 months	6 months	6 months[c]
>6 months	6 months	Chronic[c]

[a] In Japan, if there are no phase 2 clinical trials of equivalent duration to the proposed phase 3 trials, conduct of longer duration toxicity studies is recommended as given in Table 14.2.

[b] In the United States, as an alternative to 2-week studies, single dose toxicity studies with extended examinations can support single dose human trials.

[c] Data from 6 months of administration in nonrodents should be available before initiation of clinical trials longer than 3 months. Alternatively, if applicable, data from a 9-month nonrodent study should be available before clinical treatment duration exceeds that which is supported by the available toxicology studies.

Source: Taken from the ICH M3 guideline. Nonclinical Safety Studies for the Conduct of Human Clinical Trials for Pharmaceuticals, November 1997 and updated November 2000, Available at http://www.fda.gov/cder/guidance.

TABLE 14.2

Duration of Repeated Dose Toxicity Studies to Support Phase 3 Clinical Trials in the EU, and Product Marketing in All Jurisdictions[a]

	Minimum Duration of Repeated Dose Toxicity Studies	
Duration of Clinical Trials	Rodents	Nonrodents
≤2 weeks	1 month	1 month
≤1 month	3 months	3 months
≤3 months	6 months	3 months
>3 months	6 months	Chronic

[a] The above table also reflects the marketing recommendations in all three ICH regions except that a chronic nonrodent study is recommended for clinical use > 1 month.

Source: Taken from the ICH M3 guideline. Nonclinical Safety Studies for the Conduct of Human Clinical Trials for Pharmaceuticals, November 1997 and updated November 2000, Available at http://www.fda.gov/cder/guidance.

have been available for some time.[14,15] In general, for initial repeated dose studies, protocols require the use of three dose groups plus a control, and a minimum of 10 rodents and 3 to 4 nonrodents per sex per group. Doses must be selected that will allow for the identification of toxic effects at the highest dose as well as a no-effect level at the middle or lowest dose.

Usual experimental procedures include the determination of body weights and food consumption on at least a weekly basis, evaluation of hematology, and blood chemistry parameters during the treatment period, ophthalmoscopic examinations, and the recording of macroscopic examinations at necropsy, and the determination of organ weights. Electrocardiographic studies are needed in nonrodent studies. A complete histopathological examination of tissues from animals is required. In rodent studies, this can take the form of examination of all high dose and control animals and also target organs at the two lower doses. In nonrodent studies, it is typical to examine tissues from all animals in the study. Toxicokinetic studies for the same studies are required (see Section 14.10).

14.7.3 Mutagenicity

Mutagenicity studies are designed to evaluate the potential for a chemical agent to have an effect on DNA and chromosomes. There is a natural ability of mammalian cell systems to repair damage to DNA. However, when the repair mechanism fails or there is a defect in the actual repair, then the result can either be a mutation in the gene or there can be damage to the chromosome. Research in mutagenicity escalated in 1970s when it was thought that these studies may be replacements for the long and costly carcinogenicity studies, since some carcinogens act through genetic mechanisms. This goal was never realized because not all carcinogens act through a DNA-interactive mechanism. Nevertheless, mutagenicity studies provide useful indications of the ability of a drug to alter genetic material that may later be manifested in studies of carcinogenic or teratogenic potential. Mutagenicity studies are relatively inexpensive and may also serve, early in the drug development process, to assure drug developers and regulators that no obvious risk of such adverse effects exists, albeit knowing that more definitive studies to evaluate teratogenic and carcinogenic effects will not come until later.

An exhaustive review of the theory and mechanistic considerations of a mutagenicity evaluation are too lengthy for this discussion. ICH guidelines are available and include general principles and specifics related to the core battery of studies required.[16,17] Initial INDs are generally required to have two *in vitro* assays submitted, with the *in vivo* assay submitted prior to phase 2.

Positive results in one or more mutagenicity assays do not necessarily translate into human risks. FDA has issued a guidance document on approaches and considerations when positive mutagenicity studies are encountered.[18] Among these considerations are the performances of additional studies. Mechanistic studies may show that such responses would not occur in the human cell population, or the concentrations at which positive responses occurred may far exceed any concentration of drug that may occur in the clinical setting. Many marketed drugs have produced some type of positive response in these studies and yet it has been concluded that no human risk is present or the potential risk is not known. See Chapter 15 for more details.

14.7.4 Safety Pharmacology

Issues of cardiovascular safety have brought safety pharmacology to the forefront, especially with issues like prolongation of the QTc interval. In the United States, IND's must routinely include the core battery of cardiovascular, CNS, and pulmonary function studies. The reader is referred to Chapter 13 for a thorough discussion of this important topic.

14.8 Supporting Phase 2 Studies and Beyond

Once a candidate material reaches phase 2 and beyond, the nonclinical program continues, but is far more likely to deviate from some standard approach. As shown in Tables 14.1 and 14.2, the ultimate clinical regimen dictates the ultimate duration of the repeated dose animal studies. For example, a diagnostic agent or a drug with a 3 to 4 day regimen (as might be the case for drugs that may treat acute disease or trauma such as stroke or myocardial infarction that are handled in the ICU) may require little in the way of additional repeated dose toxicity studies. On the other hand, chronic therapy drugs, such as antidiabetic or antihypertensive agents, may require chronic and carcinogenicity studies. However, it is likely that the nonclinical database will contain additional studies that are designed to follow-up on identified toxicities and may include mechanistic studies that show that the effect is not likely to occur in humans because of the species-specific nature of the response.

The following types of studies are important to the continued development of most drugs.

14.8.1 Chronic Studies

In many cases, 90-day studies are needed to meet interim clinical study needs, but eventually chronic studies will be needed either to initiate phase 3 studies of greater than 3 months duration or for marketing. Chronic studies may also be needed in the case where phase 3 studies can be less than 90 days duration to show efficacy, but the disease is considered chronic in the marketplace and the drug may be used periodically. Chronic studies are generally 6 months. The U.S. FDA requires nonrodent studies to be 9 months in duration, and has required them to be 12 months if the drug is a "first in class" entity. The ICH guidelines also describe situations where studies of 9 to 12 months duration in a nonrodent species may be necessary. DeGeorge et al.[19] also concluded that there was enough data available to support a harmonized 9-month duration.

Protocols for these studies are similar to those for studies of shorter duration, except that a minimum of 15 to 20 rodents per group and 4 nonrodents per sex per group are required. All of the previously described in-life and postmortem studies are required. Toxicokinetic measurements are required at time points up to and including the termination of the study.

14.8.2 Reproduction Studies

Some of the earliest guidelines were issued by the U.S. FDA.[3] Several ICH guidelines now cover the performance of these studies, including amendments in 1995 to address possible effects on male reproduction.[20,21]

All of these guidelines have identified three phases of the entire reproductive process. In the first phase (historically referred to as Segment I study, and now under ICH as Stage A), rats are treated for a period of time prior to mating sufficient to cover one or more cycles of gametogenesis (usually 60 days in males and 14 days in females, and treatment continues up to approximately day 6 of gestation, sufficient to cover the period of implantation). This evaluates the effect of the new drug on fertility and early implantation. The second phase (historically Segment II, now ICH Stage B) is now termed the developmental toxicity phase (previously called teratology studies, and still are in many circles) and is done in both rats and rabbits. Treatment of the females occurs from day 6 and continues through day 16 (rats) and day 20 (rabbits). The fetuses are taken by cesarian section and are evaluated for skeletal

and visceral malformations. The third phase (Segment III or ICH Stage C) has female rats treated beginning in late gestation and through parturition and lactation. The offspring are allowed to deliver naturally, and behavioral and developmental assessments of the offspring are made.

The period in the drug development process at which results of these studies are required varies somewhat from country to country, but recommendations are made in the ICH guidelines. In most instances, Stage B studies are required prior to phase 2 clinical studies, the Stage A study is required prior to phase 3, and the Stage C study is needed for the NDA. Circumstances for an individual drug or requirements of certain countries may warrant a modification of this schedule.

14.8.3 Carcinogenicity Studies

Carcinogenicity studies are a time-consuming (up to 3 years to complete) and expensive (can be in excess of 1 million dollars each) proposition and are typically carried out in rats and mice. Their purpose is to determine if the drug possesses the capability to initiate or promote the development of tumors. The application of the results of these studies to the human safety has been debated for many years. In many instances over the past two decades, mechanistic studies have shown that positive responses in these rodent models do not have specific relevance for humans, and drugs have been approved on the basis of these explanations. While the scientific debate about relevance of these studies continues, they remain required by regulations. Positive responses without adequate explanation or safety margin can result in nonapproval of the product.

ICH guidelines have been issued that address the various critical elements as to when studies are needed and appropriate design characteristics. Carcinogenicity studies are generally needed when clinical treatment is longer than 3 months. There are obvious exceptions and the reader is referred to the ICH guideline that addresses the need for these studies.[22,23] Other features of the new drug may mandate carcinogenicity testing (irrespective of clinical duration), such as structure–activity similarities to known carcinogens, evidence of preneoplastic lesions in repeated dose nonclinical studies, or long-term tissue sequestration of the drug in a particular tissue. Other ICH guidelines address the complex issue of the selection of doses for these studies.[24,25] This guideline allows for flexibility in dose selection, particularly the highest dose, where the use of the maximum tolerated dose had been the primary basis. The current ICH guideline allows an option for high dose to be based on toxicokinetic data, granting that a high dose does not have to be higher than that which produced an AUC 25-fold over that seen in clinically relevant doses at steady state. An amendment to this guideline further states that the highest dose in a carcinogenicity study need not exceed 1500 mg/kg/day (1) when there is no evidence of genotoxicity; and (2) when the maximum recommended human dose is no more than 500 mg/day.[25] FDA has issued guidelines for the formatting and analysis of tumor data.[26]

A review of carcinogenicity testing procedures addresses the many factors that should be considered in a carcinogenicity program.[27]

Current protocols for carcinogenicity studies have changed little since the early 1980s. These protocols were later published as OECD guidelines.[14] There was a lot of thought that mice, especially in the 2-year study, were unusually sensitive to certain classes of chlorinated hydrocarbons. The ICH guideline allows for the option of using transgenic mice when the study is of 6 months duration.

There are three transgenic models that are primarily used. The p53 model is used exclusively for compounds that have been shown to be positive in one or more mutagenicity studies and are considered to be genotoxic. The Hras2 model is used for nongenotoxic

compounds that are administered by routes other than topical. The Tg.AC model is for topically applied drugs, and involves the topical application of the agent and uses the promotion of papillomas as the endpoint. Several of the first compounds that utilized this model produced a positive response. Controversy centers around the presence or absence of inflammation or hyperplasia and either of their roles in the papilloma formation.

The relevance of carcinogenicity testing is a continuing source of differing opinions, and some thoughtful insights have been proposed.[28,29]

14.9 Immunotoxicity

The immune system, by which the body is protected from infectious processes, is a very complex system, involving various cellular and cytokine components. These components can, when stimulated by outside pathogenic agents, respond in a rapid and complex fashion to assist the body in overcoming infectious processes. In recent years, the toxicology community has become very aware that it is possible for drugs and other chemicals to interfere with this critical system and weaken the body's ability to mount a necessary immunologic response, usually referred to as immunosuppression. Another possibility is immunogenicity, where the outside agent brings about an immunologic reaction, such as the body response to a foreign protein material, which can create an antibody response. Hypersensitivity reactions can occur where the body becomes "sensitized" to a material and will respond in a potentially harmful manner when exposed to subsequent doses of the material. A complete discussion of all of these complex processes would be too lengthy to be presented here. Suffice it to say that evaluation of a drug's potential effects on the immune system is now an integral part of any nonclinical program.

Immunosuppression can usually first be detected in the repeated dose toxicology studies through effects on globulin levels, hematological assessments such as effects on the white blood cell parameters, and organ weights and histopathological evaluation of tissues primarily associated with the immune system such as spleen, thymus, lymph node (including those that drain the primary sites of drug exposure), and bone marrow. If effects are noted, a variety of immune system functional assays are available to follow-up such findings.

Immunogenicity can be identified by identifying potential antibodies in plasma. These usually involve ELISA methods. T-cell-dependent antibody responses can be evaluated, and plaque assays involving IgM antibody responses are available.

Sensitization can be evaluated in the well-known delayed contact hypersensitivity assay or the mouse local lymph node assay.

A variety of guidance documents are available that discuss the various facets of this area of interest, including EMEA.[30] The FDA has issued a guidance document that has an excellent discussion of the background on this topic as well as recommended approaches for assessment.[31] Dean[32] has provided an insightful assessment of the history of this field, direct experience with a variety of agents that have been evaluated, and a discussion of how the various regulatory agencies in the three jurisdictions vary in their approach to this area of assessment.

14.10 Toxicokinetics

It is increasingly important that plasma concentrations, particularly C_{max} and AUC, are able to be related to dose at which toxicity is seen as well as doses at which no toxicity is seen.

This exposure relationship is frequently more important in establishing human safety margins, as dose alone may be subject to a variety of differences between species such as absolute bioavailability, distribution, and excretion. This aspect, now commonly referred to as "toxicokinetics," has been outlined in an ICH guideline.[6] This guideline specifies minimum requirements in terms of number of time points examined, number of animals per time point, and the requirements for calculation of various pharmacokinetic parameters such as C_{max}, AUC. These will become important for comparison with human data as it becomes available later.

14.11 Other Studies

No study can address the multiplicity of other studies that may have to be conducted in order to successfully bring a new drug to the market. Many regulatory agencies, including the FDA, require studies in pediatric populations, and frequent studies in young animals may be necessary to assess potential safety concerns. Parenteral drugs require assessment of the potential for local irritating effects. Topically applied drugs, for which either the active ingredient or the formulation absorb ultraviolet light in the 200 to 800 nm range, must be evaluated for phototoxicity and photosensitization, and may require photocarcinogenicity testing. As mentioned previously, mechanistic studies may be necessary to explain unusual or species-specific effects of potential new drugs.

It suffices to say that the job of the toxicologist in pharmaceutical development programs is a constant challenge, when all of the points brought out in this chapter are considered. Each candidate has its own unique characteristics that require thoughtful attention, and unanticipated toxicities are an ever-present occurrence. And for those successful candidates that make it through the program, the toxicologist has the responsibility to summarize the information for the final application to a regulatory agency and, perhaps, have the satisfaction of seeing the product approved.

References

1. MacGregor, J.T., SNPs and chips: genomic data in safety evaluation and risk assessment, *Toxicol. Sci.*, 73, 207, 2003.
2. Cunningham, M.L., Workshop overview: use of genomic data in risk assessment, *Toxicol. Sci.*, 73, 209, 2003.
3. Goldenthal, E.I., Current Views on Safety Evaluation of Drugs, FDA Papers, May, 1968.
4. ICH, Q3A, Impurities in New Drug Substances, Issued February 10, 2003, Available at http://www.fda.gov/cder/guidance.
5. ICH, S3A, Toxicokinetics: The Assessment of Systemic Exposure in Toxicity Studies, March 1995, Available at http://www.fda.gov/cder/guidance.
6. ICH, M3, Maintenance of Nonclinical Safety Studies for the Conduct of Human Clinical Trials for Pharmaceuticals, November 1997 and updated November 2000, Available at http://www.fda.gov/cder/guidance.
7. FDA, Guidance for Industry: Single Dose Acute Toxicity Testing for Pharmaceuticals, August 1996, Available at http://www.fda.gov/cder/guidance.
8. FDA, Guidance for Industry, Investigators and Reviewers; Exploratory IND Studies, April 2005, Available at http://www.fda.gov/cder.

9. ICH, S6, Preclinical Safety Evaluation of Biotechnology-Derived Pharmaceuticals, November 1997, Available at http://www.fda.gov/cder/guidance.
10. Terrell, T.G. and Green, J.D., Issues with biotechnology products in toxicologic pathology, *Toxicol. Pathol.*, 22, 187–193, 1994.
11. FDA, Guidance for Industry: Content and Format if Investigational New Drug Applications (INDs) for Phase 1 Studies of Drugs, Including Well-Characterized, Therapeutic, Biotechnology-Derived Products, November 1995, Available at http://www.fda.gov/ cder/guidance.
12. FDA, Guidance for Industry and Reviewers: Estimating the Safe Starting Dose in Clinical Trials for Therapeutics in Adult Healthy Volunteers, November 2002, Available at http://www.fda.gov/cder/guidance.
13. DeGeorge, J.J., Ahn, C., Andrews, P.A., Brower, M., Giorgio, D.W., Goheer, M.A., Lee-Ham, D.Y., McGuinn, W.D., Schmidt, W., Sun, C.J., and Tripathi, S.C., Regulatory considerations for preclinical development of anticancer drugs, *Cancer Chemother. Pharmacol.*, 41, 173, 1998.
14. Organisation for Economic Cooperation and Development, OECD Guidelines for the Testing of Chemicals, Paris, 1993.
15. Japanese Ministry of Health and Welfare. Japanese Guidelines for Nonclinical Studies of Drugs Manual 1995, Yakuji Nippo, Ltd.
16. ICH S2A, Specific Aspects of Regulatory Genotoxicity Tests for Pharmaceuticals, http://www.fda.gov/cder/guidance.
17. ICH S2B, Genotoxicity: A Standard Battery for Genotoxicity Testing of Pharmaceuticals, http://www.fda.gov/cder/guidance.
18. FDA, Guidance for Industry: Recommended Approaches to the Integration of Genetic Toxicology Study Results, November 2004, Available at http://www.fda.gov/cder/guidance.
19. DeGeorge, J.J., Meyers, L.L., Takahashi, M., and Contrera, J.F., The duration of nonrodent toxicity studies for pharmaceuticals, *Toxicol. Sci.*, 49, 143–155, 1999.
20. ICH S5A, Detection of Toxicity to Reproduction for Medicinal Products, Issued September 1994, Available at http://www.fda.gov/cder/guidance.
21. ICH S5B, Detection of Toxicity to Reproduction for Medicinal Products: Addendum on Toxicity to Male Fertility, Available at http://www.fda.gov/cder/guidance.
22. ICH S1A, The Need for Long-Term Rodent Carcinogenicity Studies of Pharmaceuticals, March 1996, Available at http://www.fda.gov/cder/guidance.
23. ICH S1B, Testing for Carcinogenicity of Pharmaceuticals, Issued February 1998, Available at http://www.fda.gov/cder/guidance.
24. ICH S1C, Dose Selection for Carcinogenicity Studies of Pharmaceuticals, Issued February 1998, Available at http://www.fda.gov/cder/guidance.
25. ICH S1C (r), Guidance on Dose Selection for Carcinogenicity Studies of Pharmaceuticals: Addendum on a Limit Dose and Related Notes, Issued November 1997, Available at http://www.fda.gov/cder/guidance.
26. FDA, Statistical Aspects of the Design, Analysis and Interpretation of Chronic Rodent Carcinogenicity Studies of Pharmaceuticals, January 2003, Available at http://www.fda.gov/cder/guidance.
27. Reno, F.E., Carcinogenicity studies, in *Comprehensive Toxicology*, *Vol. 2, Toxicological Testing and Evaluation*, Williams, P.D. and Hottendorf, G.H., Eds., Elsevier, London, 1977, pp. 121–131.
28. Cohen, S.M., Klaunig, J., Meek, M.E., Hill, R.N., Pastoor, T., Lehman-McKeeman, L., Bucher, J., Longfellow, D.G., Seed, J., Dellarco, V., Fenner-Crisp, P., and Patton, D., Evaluating the human relevance of chemically induced animal tumors, *Toxicol. Sci.*, 78, 181–186, 2004.
29. Cohen, S.M., Human carcinogenic risk evaluation: an alternative approach to the two-year rodent bioassay, *Toxicol. Sci.*, 80, 225–229, 2004.
30. EMEA, Note for Guidance on Repeated Dose Toxicity, CPMP/SWP/1042/99.
31. FDA, Guidance for Industry: Immunotoxicology Evaluation of Investigational New Drugs, October 2002, Available at http://www.fda.gov/cder/guidance.
32. Dean, J.H., A brief history of immunotoxicology and a review of the pharmaceutical guidelines, *Int. J. Toxicol.*, 23, 83–90, 2004.

15

Preclinical Genotoxicity Testing — Past, Present, and Future

Richard H.C. San

CONTENTS

15.1 Historical Background and Current Regulatory Requirements

Following the first demonstration of the deleterious effect of radiation and ultimately that of chemicals on genetic material, numerous test systems have been used to study the induction of DNA damage, chromosomal aberrations, and mutations. This broad spectrum of activity resulted in the birth of genetic toxicology. The test organisms include prokaryotes (e.g., bacteria, fungi) and eukaryotes (e.g., yeast, fruit flies, plants, mammals).

The advent of these test systems was followed by the screening of chemicals and determining (for each test system and comparison among test systems) the correlation between genotoxicity and rodent carcinogenicity. This exercise led to the formulation of genotoxicity test batteries, culminating in the adoption of harmonized test guidelines for the preclinical safety testing of pharmaceuticals. Through the International Conference on Harmonization (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use, a number of ICH harmonized tripartite guidelines have been finalized.

Guideline S2A provides "Guidance on Specific Aspects of Regulatory Genotoxicity Tests for Pharmaceuticals."[1] Guideline S2B specifically addresses "Genotoxicity: A Standard Battery for Genotoxicity Testing of Pharmaceuticals."[2]

Under Guideline S2B, the following standard test battery is recommended: (1) a test for gene mutation in bacteria, (2) an *in vitro* test with cytogenetic evaluation of chromosomal damage with mammalian cells *or* an *in vitro* mouse lymphoma thymidine kinase (TK) assay, and (3) an *in vivo* test for chromosomal damage using rodent hematopoietic cells.

15.2 Bacterial Mutation Test

In the bacterial mutation test, the mutagenic potential of a pharmaceutical and its metabolites is evaluated by measuring and quantifying its ability to induce reverse mutations at selected loci of *Salmonella typhimurium* or *Escherichia coli* in the presence and absence of metabolic activation. This test system has been shown to detect a diverse group of chemical mutagens.[3,4] The technical details of this test have been reported in the literature.[5–7]

15.3 *In Vitro* Cytogenetic Test

In the *in vitro* cytogenetic test, the clastogenic or chromosome breakage potential of a pharmaceutical and its metabolites is evaluated by measuring and quantifying its ability to induce chromosome aberrations in culture using an established cell line or a primary cell source. The use of cell cultures as a test system has been demonstrated to be an effective method of detection of chemical clastogens.[8] Induction of chromosome breakage *in vitro* is an indicator that the test article is potentially genotoxic. The cell types routinely used in these assays are Chinese hamster ovary (CHO) cells, Chinese hamster lung (CHL) cells, and human peripheral blood lymphocytes (HPBL).

15.4 *In Vitro* Mammalian Cell Gene Mutation Test (Using Mouse Lymphoma L5178Y $TK^{+/-}$ Cells)

In the *in vitro* mouse lymphoma TK mutation assay, the mutagenic potential of a pharmaceutical or its metabolites is evaluated by quantitating forward mutations at the TK locus in L5178Y mouse lymphoma cells.[9] The assay utilizes L5178Y cells which are heterozygous at the TK locus. Potential mutagenic agents are tested for the ability to cause the $TK^{+/-} \rightarrow TK^{-/-}$ mutation. $TK^{-/-}$ mutants lack the salvage enzyme TK and can easily be detected by their resistance to lethal thymidine analogs. The selective agent of choice is trifluorothymidine (TFT).[10] In addition to the detection of point mutation, the use of L5178Y cells also permits the detection of mutants resulting from chromosomal rearrangements. Large-colony mutants are believed to have received very localized damage, possibly in the form of point mutation or small deletion within the TK locus, whereas small-colony mutants received damage to large segments of the chromosomal material along with the loss of TK activity.[11–13]

15.5 *In Vivo* Micronucleus Test

The *in vivo* micronucleus test is used for the detection of damage to chromosomes as well as the mitotic apparatus in bone marrow or peripheral blood cells of rodents. The assay system has been well standardized.[14–17] The basic features of the test system are (1) the effect of the test chemical is observed in anucleated polychromatic erythrocytes (PCEs); (2) PCEs have a relatively short lifespan, so that any micronuclei they contain must have been generated as a result of recently induced chromosome damage; (3) micronuclei are readily identifiable and their distribution is well defined; and (4) the frequency of induced micronuclei in PCEs is dependent on sampling times.

Erythroblasts in bone marrow undergo a final chromosome replication after which they divide and differentiate into PCEs. Chromosomal breaks or interference in the mitotic process that result in the lagging chromosomes during this division lead to the formation of micronuclei that are similar in appearance but much smaller than the nucleus in immature, nucleated erythrocytes. During differentiation, only the nucleus is expelled from the nucleated erythrocyte, leaving behind any micronuclei formed.

The micronucleus assay may be used not only for the detection of acute but also chronic genetic damage. In mice, chromosomal breakage in bone marrow erythroblasts produces an accumulation of micronuclei in normochromatic erythrocytes in peripheral blood and there is little, if any, selective removal of micronucleated cells from circulation. This is not the case with rats, which limits their usefulness in long-term studies using peripheral blood.

15.6 Additional Genotoxicity Tests

If the results of the tests specified in the ICH guidance documents do not indicate that the pharmaceutical under test elicited any genotoxic activity, typically no further testing for genotoxicity is required. However, if genotoxicity is manifested in any of the tests in the ICH core battery, then additional testing will usually be required. The additional genotoxicity tests may include but not be limited to one of the following tests: *in vivo* unscheduled DNA synthesis (UDS) test in rat hepatocytes, *in vitro* Syrian hamster embryo (SHE) cell transformation test, and *in vivo* (or *in vitro*) single-cell gel electrophoresis test for DNA damage (Comet assay).

15.6.1 Unscheduled DNA Synthesis Test Using Primary Rat Hepatocytes

Monitoring DNA repair synthesis is a widely used method for assessing DNA-damaging activity. As opposed to the scheduled DNA synthesis that occurs during the normal phase of semiconservative duplication of DNA in the cell cycle, DNA synthetic activity triggered by DNA damage can occur at any phase of the cell cycle and is commonly referred to as "unscheduled DNA synthesis"(UDS). Measurement of UDS can be achieved by tracking the incorporation of BrdUrd or tritiated thymidine into nuclear DNA of repairing cells, although other purine or pyrimidine precursors can also be used.[18] The incorporation of radioactively labeled purine or pyrimidine can be measured by autoradiographic or scintillation counting methods. Examination of over 100 chemical compounds representing the major groups of carcinogenic substances revealed a good correlation between the carcinogenic activity and

the capacity to elicit DNA repair synthesis in cultured mammalian cells.[19,20] A review of the published literature and suggested protocols and evaluation criteria for evaluating UDS were presented in a working group report prepared for the Gene-Tox Program, U.S. Environmental Protection Agency.[21] Recommendations for the performance of UDS assays *in vitro* and *in vivo* have been presented in subsequent papers.[22–24]

15.6.2 *In Vitro* Syrian Hamster Embryo Cell Transformation Test

Under the strategy for testing carcinogenic potential of a pharmaceutical, the ICH harmonized tripartite guideline S1B indicated that data from *in vitro* assays, such as a cell transformation assay, can be useful at the compound selection stage.[25] This has triggered an interest by regulatory agencies for the SHE cell transformation assay for a genotoxic pharmaceutical, especially if the drug is intended to be administered chronically to humans.

The SHE cell transformation assay at pH 6.7 is conducted to determine the ability of a test article to induce morphological transformation in cultured SHE cells. The SHE cell transformation assay is one of the most widely used cell transformation assays. The endpoints of this assay are related to the conversion of normal cells into preneoplastic or neoplastic cells. The assay provides a valuable tool in the process of assessment and evaluation of the carcinogenic potential of a test article.[26–28]

The SHE cell transformation assay is conducted with cultured Syrian hamster embryonic cells prepared at the gestation period of approximately 13 days. Cells from the embryos are isolated and frozen. A feeder layer of cells is seeded with x-ray-irradiated SHE cells prior to seeding the target cells. The SHE transformation assay is performed at pH 6.7. The assay is designed to allow the expression of transformed morphology of clonal SHE cells seeded onto feeder cells, after exposure to at least five concentrations of test article as well as positive and vehicle controls in the absence of a supplemental exogenous mammalian metabolic activation system. For maximum assay sensitivity, two exposure durations can be used, namely, a 24-hr and a 7-day exposure assay. Following treatment, the cells are cultured for a period of 7 days, and then fixed, stained, and evaluated for the cytotoxic effects of treatment and the induction of phenotypic transformation as measured by morphological transformation of the cells.[29–31]

15.6.3 *In Vivo* (or *In Vitro*) Single-Cell Gel Electrophoresis Test for DNA Damage (Comet Assay)

The single-cell gel electrophoresis test for DNA damage, commonly known as the Comet assay, is a sensitive technique unique for its ability to measure DNA damage and repair in individual cells. The Comet assay is a relatively new technique that is increasingly being used in the testing of new pharmaceutical or chemical compounds for genotoxicity and for environmental monitoring, human biomonitoring and molecular epidemiology, and fundamental research in DNA damage and repair.[32–34]

The innovation with the Comet assay was introduced by Östling and Jonhanson[35] when they embedded cells into agarose gels to immobilize them. Then the cells were incubated in detergents and high salts to lyse the cellular membrane and break down proteins and RNA, to allow the DNA to relax. Subsequently, the cells underwent electrophoresis to enable the relaxed DNA strands to move out of the nucleus. This resulted in the formation of a "tail" on one side of the cell. Because this image resembles a comet, this technique is now universally known as the Comet assay. With a modification introduced by two independent groups, pretreatment of the cells and conducting electrophoresis at pH 13 permit the measurement of DNA single-strand breaks induced by alkaline-labile sites.[36]

The Comet assay has the ability to contribute to sensitively assess DNA damage in almost any eukaryotic cell from *in vitro* cultures or from animal organ tissues.[37,38] The *in vivo* Comet assay can be used as a supplemental assay for mechanistic or target-organ-specific toxicity.[39,40] With new modifications of the technique, DNA double-strand breaks, DNA oxidative damage, and DNA repair can be easily detected.[38] Although there are currently no ICH or OECD guidelines for the Comet assay, scientists from academia, industry, and government have recently assembled an international committee for this purpose.

15.7 Looking into the Future

The demonstration of genotoxic activity in a pharmaceutical molecule may not necessarily translate into discontinuation of development and ultimate licensing of the product to be marketed. A survey of the Physicians' Desk Reference reveals many examples of pharmaceutical products that have been shown to elicit genotoxic activity in one or more genetox assays.[41] This observation indicates that other factors (such as risk, benefit, seriousness of ailment, target patient population, dosage, and frequency of administration, among others) are taken into consideration.

Computational models are currently being used by regulatory agencies and with the pharmaceutical industry to predict the mutagenic potential of new pharmaceutical candidates.[42] However, these models rely heavily on bacterial mutagenicity data of nonpharmaceutical-type chemical molecules as the knowledge base. To what extent this selection of database has limited the ability of these programs to predict genotoxicity of pharmaceuticals remains to be seen.

With the advent of new technologies (e.g., microarrays) to study the effect of chemicals (and pharmaceuticals) on the upregulation or downregulation of gene activity, the field of toxicogenomics is quickly moving past its infancy. Although these approaches may be viewed by some as information overload, they have the potential to serve as indicators of early events occurring in cells following exposure to chemicals, and ultimately as predictors of genotoxic activity.

Despite the usefulness of genotoxicity tests for the identification of synthetic chemical compounds as potential human health hazards, it is recognized that there are many naturally occurring environmental mutagens (e.g., plant toxins) that may contribute to human diseases.[42] In addition, deficiency of some of the micronutrients required in the human diet (e.g., vitamins B_{12}, folic acid, B_6, niacin, C, or E, or iron, or zinc) appears to mimic radiation in damaging DNA by causing single- and double-strand breaks, oxidative lesions, or both.[43] In the evaluation of human health hazards, it would be appropriate to view all these contributing factors in the proper perspectives.

References

1. International Conference on Harmonisation (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use. Guidance on Specific Aspects of Regulatory Genotoxicity Tests for Pharmaceuticals. S2A document recommended for adoption at step 4 of the ICH process on July 19, 1995, Federal Register 61:18198-18202, April 24, 1996.
2. International Conference on Harmonisation (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use. Genotoxicity: A Standard Battery for Genotoxicity Testing

of Pharmaceuticals. S2B document recommended for adoption at step 4 of the ICH process on July 16, 1997, Federal Register 62:16026-16030, November 21, 1997.
3. McCann, J. and Ames, B.N., Detection of carcinogens as mutagens in the *Salmonella*/microsome test: assay of 300 chemicals: discussion, *Proc. Natl. Acad. Sci. USA*, 73, 950, 1976.
4. McCann, J., Choi, E., Yamasaki, E., and Ames, B.N., Detection of carcinogens as mutagens in the *Salmonella*/ microsome test: assay of 300 chemicals, *Proc. Natl. Acad. Sci. USA*, 72, 5135, 1975.
5. Ames, B.N., McCann, J., and Yamasaki, E., Methods for detecting carcinogens and mutagens with the *Salmonella*/ mammalian–microsome mutagenicity test, *Mutat. Res.,* 31, 347, 1975.
6. Green, M.H.L. and Muriel, W.J., Mutagen testing using trp^+ reversion in *Escherichia coli*, *Mutat. Res.,* 38, 3, 1976.
7. Maron, D.M. and Ames, B.N., Revised methods for the *Salmonella* mutagenicity test, *Mutat. Res.,* 113, 173, 1983.
8. Preston, R.J., Au, W., Bender, M.A., Brewen, J.G., Carrano, A.V., Heddle, J.A., McFee, A.F., Wolff, S., and Wassom, J.S., Mammalian *in vivo* and *in vitro* cytogenetic assays: a report of the Gene-Tox Program, *Mutat. Res.,* 87, 143, 1981.
9. Clive, D. and Spector, J.F.S., Laboratory procedure for assessing specific locus mutations at the TK locus in cultured L5178Y mouse lymphoma cells, *Mutat. Res.,* 31, 17, 1975.
10. Brown, M.M.M. and Clive, D., The utilization of trifluorothymidine as a selective agent for $TK^{-/-}$ mutants in L5178Y mouse lymphoma cells, *Mutat. Res.,* 53, 116, 1978.
11. Clive, D., Johnson, K.O., Spector, J.F.S., Batson, A.G., and Brown, M.M.M., Validation and characterization of the L5178Y $TK^{+/-}$ mouse lymphoma mutagen assay system, *Mutat. Res.,* 59, 61, 1979.
12. DeMarini, D.M., Brockman, H.E., de Serres, F.J., Evans, H.H., Stankowski Jr, L.F., and Hsie, A.W., Specific-locus mutations induced in eukaryotes (especially mammalian cells) by radiation and chemicals: a perspective, *Mutat. Res.,* 220, 11, 1989.
13. Hozier, J., Sawyer, J., Moore, M., Howard, B., and Clive, D., Cytogenetic analysis of the L5178Y/$TK^{+/-}$? $TK^{-/-}$ mouse lymphoma mutagenesis assay system, *Mutat. Res.,* 84, 169, 1981.
14. Heddle, J.A., A rapid in vivo test for chromosomal damage, *Mutat. Res.,* 18, 187, 1973.
15. Matter, B.E. and Grauwiler, J., Micronuclei in bone marrow cells. A simple in vivo model for the evaluation of drug-induced chromosomal aberrations, *Mutat. Res.,* 23, 239, 1974.
16. Heddle, J.A., Hite, M., Kirkhart, B., Mavournin, K., Macgregor, J.T., Newell, G.W., and Salamone, M., The induction of micronuclei as a measure of genotoxicity. A report of the U.S. Environmental Protection Agency Gene-Tox Program, *Mutat. Res.,* 123, 61, 1983.
17. Hayashi, M., Morita, T., Kodama, Y., Sofuni, T., and Ishidate, Jr, M., The micronucleus assay with peripheral blood reticulocytes using acridine-coated slides, *Mutat. Res.,* 245, 245, 1990.
18. Cleaver, J.E., DNA repair with purines and pyrimidines in radiation- and carcinogen-damaged normal and *Xeroderma pigmentosum* human cells, *Cancer Res.,* 33, 362, 1972.
19. San, R.H.C. and Stich, H.F., DNA repair synthesis of cultured human cells as a rapid bioassay for chemical carcinogens, *Int. J. Cancer,* 16, 284, 1975.
20. Stich, H.F. and San, R.H.C., Reduced DNA repair synthesis in *Xeroderma pigmentosum* cells exposed to the oncogenic 4-nitroquinoline 1-oxide and 4-hydroxayminoquinoline 1-oxide, *Mutat. Res.,* 13, 279, 1971.
21. Mitchell, A.D., Casciano, D.A., Meltz, M.L., Robinson, D.E., San, R.H.C., Williams, G.M., and Von Halle, E.S., Unscheduled DNA synthesis tests: a report of the Gene-Tox program, *Mutat. Res.,* 123, 363, 1983.
22. Butterworth, B.E., Ashby, J., Bermudez, E., Casciano, D., Mirsalis, J., Probst, G., and Williams, G., A protocol and guide for the in vitro rat hepatocyte DNA-repair assay, *Mutat. Res.,* 189, 113, 1987.
23. Butterworth, B.E., Ashby, J., Bermudez, E., Casciano, D., Mirsalis, J., Probst, G., and Williams, G., A protocol and guide for the in vivo rat hepatocyte DNA-repair assay, *Mutat. Res.,* 189, 123, 1987.
24. Madle, S., Dean, S.W., Andrae, U., Brambilla, G., Burlinson, B., Doolittle, D.J., Furihata, C., Hertner, T., McQueen, C.A., and Mori, H., Recommendations for the performance of UDS tests in vitro and in vivo, *Mutat. Res.,* 312, 263, 1994.
25. International Conference on Harmonisation (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH Harmonised Tripartite Guideline. Testing for

Carcinogenicity of Pharmaceuticals. S1B document recommended for adoption at step 4 of the ICH process on 16 July 1997, Federal Register 63:8983, 23 February 1998.
26. Isfort, R.J., Cody, D.B., Doerson, C., Kerckaert, G.A., and LeBoeuf, R.A., Alterations in cellular differentiation, mitogenesis, cytoskeleton and growth characteristics during Syrian hamster embryo cell multistep in vitro transformation, *Int. J. Cancer,* 59, 114, 1994.
27. Kerckaert, G.A., Isfort, R.J., Carr, G.J., Aardema, M.J., and LeBoeuf, R.A., A comprehensive protocol for conducting the Syrian hamster embryo cell transformation assay at pH 6.70, *Mutat. Res.,* 356, 65, 1996.
28. Engelhardt, G., Schwind, K.-R., and Mußler, B., The testing of chemicals in the Syrian hamster embryo (SHE) cell transformation assay for assessment of carcinogenic potential, *Toxicology In Vitro*, 8, 213, 2004.
29. LeBoeuf, R.A., and Kerckaert, G., Enhanced morphological transformation of early passage Syrian hamster embryo cells cultured in medium with a reduced bicarbonate concentration and pH, *Carcinogenesis*, 8, 680, 1987.
30. LeBoeuf, R.A., Kerckaert, G.A., Aardema, M.J., and Gibson, D.P., Multistage neoplastic transformation of Syrian hamster embryo cells cultured at pH 6.70, *Cancer Res.,* 50, 3722, 1990.
31. Kerckaert, G.A., Isfort, R.J., Carr, G.J., Aardema, M.J., and LeBoeuf, R.A., A comprehensive protocol for conducting the Syrian hamster embryo cell transformation assay at pH 6.70, *Mutat. Res.,* 356, 65, 1996.
32. Singh, N.P., McCoy, M.T., Tice, R.R., and Schneider, E.L., A simple technique for quantitation of low levels of DNA damage in individual cells, *Exp. Cell Res.,* 175, 184, 1988.
33. Tice, R.R., Agurell, E., Anderson, D., Burilinson, B., Hartmann, A., Kobayashi, H., Miyamae, Y., Rojas, E., Ryu, J.C., and Sasaki, Y.F., Single cell gel/comet assay guidelines for in vitro and in vivo genetic toxicology testing, *Environ. Mol. Mutagen.,* 35, 206, 2000.
34. Collins, A.R., The comet assay for DNA damage and repair: principles, applications, and limitations, *Mol. Biotechnol.,* 26, 249, 2004.
35. Östling, O. and Johanson, K.J., Microelectrophoretic study of radiation-induced DNA damages in individual mammalian cells, *Biochem. Biophys. Res. Commun.,* 123, 291, 1984.
36. Olive, P.L., Banáth, J.P., and Durand, R.E., Heterogeneity in radiation-induced DNA damage and repair in tumor and normal cells measured using the "comet" assay, *Radiat. Res.,* 122, 86, 1990.
37. Hartamann, A., Agurell, D., Beevers, C., Brendler-Schwaab, S., Burlinson, B., Clay, P., Collins, A., Smith, A., Speir, G., Thybaud, V., and Tice, R.R., Recommendations for conducting the in vivo alkaline comet assay, *Mutagenesis,* 18, 45, 2003.
38. Hartamann, A., Schumacher, M., Plappert-Helbig, U., Lowe, P., Suter, W., and Mueller, L., Use of the alkaline in vivo comet assay for mechanistic genotoxicity investigation, *Mutagenesis,* 19, 51, 2004.
39. Sasaki, Y.F., Kawaguchi, S., Kamaya, A., Ohshita, M., Kabasawa, K., Iwama, K., Taniguchi, K., and Tsuda, S., The comet assay with 8 mouse organs: results with 39 currently used food additive, *Mutat. Res.,* 519, 103, 2002.
40. Sekihashi, K., Yamamoto, A., Matsummura, Y., Ueno, S., Watanabe-Akanuma, M., Kassie, F., Knasmuller, S., Tsuda, S., and Sasaki, Y.F., Comparative investigation of multiple organs of mice and rats in the comet assay, *Mutat. Res.,* 517, 53, 2002.
41. Snyder, R.D., Pearl, G.S., Mandakas, G., Choy, W.N., Goodsaid, F., and Rosenblum, I.Y., Assessment of the sensitivity of the computational programs DEREK, TOPKAT, and MCASE in the prediction of the genotoxicity of pharmaceutical molecules, *Environ. Mol. Mutagen.,* 43, 143, 2004.
42. Ames, B.N. and Gold, L.S., Cancer prevention strategies greatly exaggerate risks *C&EN News*, American Chemical Society, Washington, D.C., January 7, 1991, 28.
43. Ames, B.N., DNA damage from micronutrient deficiencies is likely to be a major cause of cancer, *Mutat. Res.,* 475, 7, 2001.

Section III

Standard Drug Development Issues: Updated

16

The Need for Animals in Biomedical Research

Charles G. Smith

CONTENTS

Few topics in biomedical research elicit the emotional response and frequent overreaction as does the subject of the use of live animals in the drug discovery and development process. In recent years, a worldwide movement of highly motivated activists has adopted, as one of its causes, the prohibition of the use of intact animals in research. Common arguments put forward to support this philosophy include:

1. Animals have contributed very little to medical discoveries in the past and may actually slow down the process of new drug discovery.
2. Many animal experiments are repetitive and wasteful.
3. Some, if not many, of the experiments are performed in an inhumane or torturous manner.
4. Much, if not all, of the information that is gleaned from animals can be obtained by the use of lower forms of life such as microorganisms and mammalian cells grown or maintained *in vitro*.
5. Computer simulation can be used to replace animals much more than it is now.

All of the above are either fallacious or grossly overstated, for reasons that will be discussed further in this chapter. To prohibit the use of live animals in biomedical research would so drastically reduce the flow of new medications into the hands of the physicians of this country and the world as to represent, in my view, an unforgivable medical disaster for generations to come.

Having said the above, a comment must immediately be made that it is incumbent on all scientists to assure that the animals used in the laboratory will be humanely treated, adequately housed, and subjected to the least pain that is necessary to achieve the objectives of well-thought-through experiments. All such investigations should be carried out in accordance with the guidelines promulgated by the National Institutes of Health, appropriate professional societies and the Department of Agriculture. The misguided sentiments of those individuals who believe that rats, dogs, rabbits, monkeys, etc., have equal rights with human beings cannot, in my opinion, be influenced by facts or reason, as their objective certainly seems to be the prevention of experimenting on intact animals under any and all circumstances. Those of us who have dedicated our lives to biomedical

research with the objective of providing new medications to alleviate human suffering and prolong human life are also dedicated to the principle that human health and welfare must be put before that of lower animals, since it is impossible to advance the cause of drug discovery and development without the use of such creatures. The above-mentioned arguments put forth by some animal rights activists will be discussed in the following paragraphs.

The history of medical research leaves no possible doubt as to the critical importance of the use of animals to discover new medications and to develop surgical and diagnostic procedures. How can anyone deny the import of the experiments of Banting and Best in the early 1920s with the hormone insulin? Can any objective person really argue that insulin should not have been available to treat diabetic patients over the last 80 years because live animals were used to discover and develop it? In fact, the hormone could not have been discovered many years ago without using animals, because today's highly sophisticated technology that permits scientists to identify tiny amounts of proteins was not available at that time. Indeed, a major catalyst for rapid development of the biotechnology field is the knowledge base that exists because of animal experimentation through which the function of hormones was demonstrated. In addition to insulin, the discovery in animals, many years ago, of the actions of steroid hormones opened tremendous horizons for the treatment of various diseases, particularly severe inflammatory diseases (such as rheumatoid arthritis and asthma), that could not have been possible without having studied their roles in the intact animal. Years ago, the only source for most of the hormones was animal tissue. The only method to demonstrate the very existence, function, and activity of hormones was to surgically remove the organs that produce the substances, observe the negative effects on the intact animal resulting from organ removal, and show the ability to reconstitute normal, or near normal, function and health upon injecting the appropriate organ extract. How can anyone really prefer that such experiments had never been done? To follow such reasoning would have denied to society the magnificent benefits of insulin, thyroid hormones, growth hormones and steroid hormones (anti-inflammatory and sex hormones), to name a few, over a period of decades.

It does not require a great deal of knowledge or experience to realize that the whole discipline of surgery could not have reached its present state of sophistication and accomplishment if it were not for experimental surgery performed on animals. If such experiments cannot be done (as surely they cannot) directly on human beings, how can we know which area of the brain, for example, can be removed in diseased or injured patients, without the risk of inflicting severe, permanent damage by virtue of the surgical procedure *per se*? The answer, clearly, is that such experiments must be performed in live animals, anesthetized to be sure, and their recovery and postsurgical function monitored. A very interesting example of the development of a new surgical procedure for the treatment of severe Parkinson's disease has been reported recently.[1,2] The same argument holds for many other surgical procedures. One of the most difficult ethical dilemmas in the animal experimentation arena, in my opinion, is the question of studying shattering traumatic and gun shot wounds. Again, one cannot develop medical or surgical procedures for dealing with such wounds by simply responding to those human patients brought into an emergency room or a battle field surgery upon whom a traumatic or gun shot injury has already been inflicted. At the same time, no normal person takes the idea of shooting a pellet into an animal lightly. The same considerations apply in the burn wound field. Even in these particularly difficult and soul-searching arenas, one cannot accept an argument that could be considered valid that would ban experimental surgery on animals.

The need for animal models of human disease to search for new therapeutic agents is also mandatory to be able to move forward with new drug discovery and development. Computers are widely used in biomedical research today, as are isolated tissues, animal

cells, enzymes, and microorganisms, but no combination of these can substitute for studies in the intact animal. Consider, in explanation of this statement, the example of a muscle cell in the myocardium of the heart. When these muscle cells begin to weaken, which will lead to heart failure, medication to stimulate their force of contraction is needed. Consider further the availability of a new drug candidate that was discovered using isolated heart muscles or cells in a test tube and which has been shown to stimulate these cells and muscle fibers *in vitro*. To deliver this drug candidate to the myocardial cells in the heart of the patient, the compound must be put into the bloodstream, either by injection or by mouth, in which it will be delivered throughout the body to all tissues, including the heart. When the patient swallows a drug, it enters the stomach where it is immediately exposed to enzymes and acid, as nature has designed the process, in an attempt to "digest" the foreign substance. The drug, if it survives the acid milieu and initial enzyme attack, then passes into the small intestine, where the pH of the environment shifts from acidic to basic and digestive juices from the pancreas are secreted to further attempt to convert the drug molecule into something more recognizable by the body. If the drug survives these barriers, it will next (1) pass through the intestinal epithelium and be carried to the liver, (2) be absorbed into the fat-absorbing system (lacteals), or (3) be excreted in the feces. Drug molecules that are absorbed into the bloodstream pass through the liver, the most highly active metabolic organ in the body. One of the very important functions of the liver is to alter or to detoxify foreign substances coming to it so that they can be excreted from the body via the kidneys or the intestine. Many drugs undergo some sort of metabolism during their course of transit from the mouth into the bloodstream and through the liver, lungs, and other organs. In the last analysis, the unchanged original drug, or its metabolites (some of which may represent the actual "active" molecules in the body), must be able to reach and penetrate the heart cells, so as to carry out the drug's biochemical reactions within the target tissue that, in turn, will ultimately increase the force of contraction of the heart.

It is absolutely impossible, no matter what any well-meaning but misguided animal protectionist may claim, to simulate the enormously complicated scenario that occurs in an animal body with any *in vitro* system, combination of *in vitro* systems, or computer programs, since we know only a fraction of what actually happens in drug absorption and delivery to target tissue. Table 16.1 presents an overview of this complexity.

The intricacy of the intact animal body is further illustrated by the example of the cytokine, tumor necrosis factor (TNF). Table 16.2 shows the spectrum of biologically active molecules that are induced by TNF and Table 16.3 lists the wide variety of biological effects observed in animals to which TNF has been administered. Figures 16.1 to 16.3 show the labyrinth of biochemical cross-talk within and among cells in the body.

What objective individual could believe that this extremely diverse system, ongoing in various cells in the animal body, could be replicated in test tube or computer experiments?

As a molecule passes from tissue to tissue via the blood, it is exposed to hundreds of possible metabolic sites; it can be bound to protein, fat, or other molecules in the serum; it can be sequestered in certain organs or excreted via the kidneys, the colon, or even the breath, at rates and by mechanisms that we simply do not completely understand. Faced with these facts, how can anyone state dogmatically that we can replace intact animals with computers, bacteria, cell cultures, etc.?

Pharmacologists attempt to develop animal models that will predict drug effects that one might see in the human being, and toxicologists conduct studies in animals to determine which organs are the most sensitive to the drug and, if an organ is toxified, whether the animal recovers when drug therapy is stopped. Such data are absolutely essential before exposing humans to a new drug candidate and there is simply no viable alternative to generating the data, other than to administer the drug candidate to an intact animal. Crude estimates of the costs to perform studies *in vitro* and in animals range from a few

TABLE 16.1

Overview of Some Facts

1. The heart recirculates the equivalent of some 7200 L (1800 gal) of blood through the body all day, every day
2. All organs, tissues, and cells are continuously perfused
3. Cells take nutrients out of and excrete waste products into the blood
4. A molecule produced by one type of cell may be essential to the function of another type of cell
5. Hormones, cytokines, and autokines permeate the trillions of cells in the body
6. How can anyone believe that the whole animal body system can be mimicked in a test tube or a computer?

TABLE 16.2

Factors Induced by Tumor Necrosis Factor

Cytokines	IL-1
	IL-6
	GM-CSF
	PDGF
	Cachectin/TNF
	TGF-β
Hormones	Corticotropin/cortisol
	Epinephrine
	Norepinephrine
	Glucagon
Eicosanoids	Platelet-activating factor
	Prostaglandins
	Leukotrienes
Acute-phase proteins	C-reactive protein
	α-1-Acid glycoprotein
	α-2-Macroglobulin
	Complement proteins factor B and C3
	α-1-Antichymotrypsin
Other	Collagenase
	Elastase
	Superoxide radicals
	Increased endothelial expression of ICAM and ELAM

Note: IL-1 = interleukin-1; GM-CSF = granulocyte-marcophage colony-simulating factor; PDGF = platelet-derived growth factor; TGF-β = transforming growth factor-β; ICAM = intercellular leukocytic adhesion molecule; ELAM = extracellular leukocytic adhesion molecule.

Source: From Tracey, K.J. and Cerami, A., *Am. J. Trop. Med. Hyg.*, 47(Pt 2), 2–7, 1992. Reproduced with permission.

thousand dollars for *ex vivo* experiments to two million dollars for carcinogenicity studies (in the rat and mouse). Shorter term *in vivo* experiments will cost thousands of dollars. If no other incentive existed for wanting to replace animal models, economic considerations would certainly drive the industry to *in vitro* test methods. All companies would rapidly abandon *in vivo* testing if it were possible to do so and still discover and develop new and safe drugs using solely *in vitro* methods and computers.

Two specific examples of the critically important value of animal models in the discovery and development of new medications follow. The first is from the cancer research field.

TABLE 16.3
Biological Effects of Tumor Necrosis Factor

Acute, high dose	Shock and death
	Tissue injury with multiple organ system failure
	Fever
	Respiratory arrest
	Pulmonary edema and ARDS
	Hemorrhagic necrosis and capillary thrombosis
	Capillary leak syndrome
	Lactic acidosis
	Hyperglycemia followed by hypoglycemia
	Catabolic hormones release
	Platelet-activating factor and eicosanoid release
Chronic, low dose	Anorexia
	Weight loss
	Anemia
	Depletion of body protein and lipid
	Hypertriglyceridemia and hyperaminoacidemia
	Increased whole-body energy expenditure
	Acute-phase protein biosynthesis
	Neutrophilia
	Catabolic hormones release
	Euthyroid-sick syndrome
Local tissue production	Fibroblast growth factor
	Angiogenesis factor
	Stimulation of growth factors (PDGF, GM-CSF, TGF)
	Induces collagenases and elastases
	Chemoattractant
	Superoxide radical release
	Localization of infection
	Increase macrophage cytotoxicity against fungus and parasites

Note: The dose ranges listed here reflect representative examples of the biological scenario from TNF activities based on *in vitro* and *in vivo* studies and do not reflect absolute for dose–response effects. See Table 16.2 for definitions.

Source: From Tracey, K.J. and Cerami, A., *Am. J. Trop. Med. Hyg.*, 47(Pt 2), 2–7, 1992.

In the late 1950s, only a few drugs were available for the treatment of acute leukemia in children. These compounds had been discovered using a combination of biochemical rationale, screening in various test tube and cell culture systems, and evaluation in animal tumor models. The median life expectancy of a child with acute lymphocytic leukemia (ALL) was, as I recall, approximately 6 to 12 months untreated and, perhaps, 1 to 2 years when treated with the drugs available at that time. If untreated, the great majority of patients with ALL were dead within 24 months of diagnosis and, even when treated, most did not survive more than a few years following diagnosis. In the late 1930s, Furth et al.[3] showed that a single, highly malignant leukemic cell was all that was required, when injected into an intact mouse, to kill the animal. The significance of this finding was that the challenge to the successful cure of cancer required killing all of the malignant cells. In the mid-1960s, Dr Howard Skipper and his colleagues[4] at the Southern Research Institute demonstrated, to a high scientific standard, that each dose of an effective chemotherapeutic agent given to tumor-bearing animals reduced the number of tumor

cells in a logarithmic mode (by a proportion of the surviving cells, rather than by an absolute number). An analogy in a cancer patient would be to say that, if we have 1 million tumor cells and give a dose of chemotherapeutic agent that will reduce the number in the body by 50%, we will kill 500,000 cells. When a second dose is given, half of the remaining cells are killed, leaving 250,000 tumor cells alive and growing. The third dose of drug will reduce the body burden to 125,000 cells, etc. At the same time that the tumor cells are being killed, the cytotoxic drugs that are used in the cancer field (the only drugs that, until recently, were available for effectively treating most cancers) are also producing toxicity in normal organs of the host and, at some point, treatment must be stopped to allow the patient to recover from drug damage to other organs that, in and by itself, can prove fatal. During the drug recovery period, the tumor cells again proliferate and repopulate the body, until the normal tissues recover and chemotherapy can again be administered. During the treatment periods, the tumor cells may become resistant to the therapy and a new treatment modality will be required if life is to be prolonged.

Based on the important findings of Furth and Skipper et al. in animals, a group of cancer chemotherapy specialists applied these principles to human beings, with truly impressive results in the treatment of ALL. Today, approximately 70 to 75% of children with acute leukemia can actually be cured. Impressive results have been obtained following the same principles in the treatment of Hodgkin's disease. There was simply no possible way, with or without computers, that biomedical scientists could have determined (1) the ability of one leukemia cell to kill an animal or (2) the so-called logarithmic kill rate of leukemic cells without using animals. Those seemingly simple experiments opened the door to curative chemotherapy of acute leukemia of childhood and certain other malignancies. It is shocking to think that any rational person could believe that those animal experiments should not have been carried out. An excellent overview of the early clinical discoveries in the chemotherapy of cancer has been published by Holland.[5]

The second example is in the area of infectious diseases. At the turn of the 19th century, the famous German chemist, Paul Ehrlich, synthesized Salvarsan, which represented the first chemotherapeutic treatment for syphilis. Compared to modern therapy, Salvarsan left a great deal to be desired. Following the discovery of penicillin in the 1950s, an all out effort was devoted to the search for new antibiotics. As a result of the blending of the sciences of microbiology, biochemistry, chemistry, pharmacology, toxicology, clinical medicine, and pharmacy, a wide spectrum of new antibiotic molecules was forthcoming. These agents have had an enormous impact on human health and survival. The research programs brought forth an array of penicillins with various specific properties (e.g., improved oral absorption, broadened antibacterial spectrum, and high bactericidal activity for certain organisms), the cephalosporin antibiotics, quinolones, macrolides, and a variety of others that are not mentioned here. Because of the availability of antibiotics over the last 50 years, many of the diseases that represented the major killers of society only 50 or 75 years ago (tuberculosis, pneumonia, septicemia, dysentery-type infections, etc.) are now curable in most of the cases. It has been said that antibiotics represent one of the key factors in the prolongation of life in the last century. Aggressive searching for new antibiotics is still required to find therapeutic agents that are active against the drug-resistant organisms that are appearing today.

No program carried out solely in test tubes could possibly have duplicated the animal models for activity or for toxicity in searching for effective antibiotics and in rank-ordering them for priority of clinical testing, then or now. When an organism infects a whole animal, the bacterium may take hold in one or more organs and is also delivered, via the bloodstream, throughout the entire body. The animal responds to this insult by attempting to kill the bacteria by calling upon its immune response system. With some organisms, the host is very successful and serious infection does not occur. With other organisms, the body

cannot defeat the intruder by itself and an infection results that may occur in one or more places in the body. There is no reliable way to imagine how this can happen or to computerize anything about the process when one does not have any facts to feed into the computer nor to make a judgment on which organisms will go on to cause devastating infection or which will be self limiting, without experimentation in animals. Clearly, such experiments cannot be done in human beings, whether volunteers or patients already ill with the disease. Although nothing was said above about discarding toxic new antibiotic candidates along the process of new drug screening and discovery, many were discarded because of significant or irreversible toxicity in animals. Again, how can anyone possibly decide which new antibiotics (or other chemotherapeutic agents) should be administered to humans and which should not without information about their effects on normal organs in the body? Obviously, whole animal toxicity studies represent the only logical approach.

Another very important use of animals is that of refining the drug development process by attempting to reduce the toxicity of drugs by chemical manipulation of the molecules. For example, certain antibiotics that were put into early clinical practice were found to be effective against organisms that cause serious infection but also were found to be unacceptably toxic to one or more organs in the body. Variants of these antibiotics were developed that are less toxic while maintaining good efficacy against the infecting bacteria. Several of these have come to represent mainstream therapies for certain severe infections in human beings. Without the animal models to compare the relative activities and toxicities (e.g., therapeutic index) of these modified antibiotics with the parent drugs, which had to be done in the complicated interactive system that exists only in an intact animal body, such advances would simply not have been possible. The selection of better therapeutic agents using the approach discussed in this paragraph is carried out in all categories of disease under investigation.

Without the availability of animal data, how would one decide on the first dose of a molecule (that has never been administered to any animal) to be given to man? A frequently used method is to give a volunteer 10% of the lowest "no-effect" dose seen in two or three animal species. How would one decide on the route of administration? Some compounds are very toxic to intact veins. How could one give reasonable assurance of safety of a new chemical never before administered to a human? The only viable option is to determine tolerance in the veins of lower animals before moving to a volunteer. Should animal activists not have to put forth their actual protocol for studying a new drug candidate in humans? How would they answer the questions raised above?

Animal rights advocates often note that drug efficacy and toxicity responses may not be the same in humans and lower animals. While this is certainly true with some drug candidates, drug effects in man are comparable to those seen in some animal species with a high percentage of compounds. In fact, the agent under development is usually tested in more than one animal species, which increases the likelihood of predictability of response in man. Correlation of activity and toxicity among species has been discussed by Smith et al.[6,7] Greaves et al.[8] present a most useful review of how animals help in the drug discovery process; they point out the requirement for animal testing prior to human experimentation in the Nuremberg Code.

Activists point to the use of computers as a mechanism for testing in fewer or no animals. The fact is that we know so little about the overall processes involved in disease and its treatment that one has little or no idea what to enter into the computer when dealing with a chemically unique drug molecule. Computers are widely used throughout the drug development process whenever possible to do so. They are especially useful when one is comparing the effects of related chemicals and in the design of new molecules. The call for replacing intact animals with lower forms of life, such as bacteria, in predicting animal efficacy or toxicity is also fallacious. Consider the fact that bacteria are one-celled organisms

that have a totally different "overcoat" around them than do human cells. That overcoat is designed, by nature, to admit or to exclude different kinds of molecules than those usually seen by human cells, because the bacteria have evolved living in a totally different environment than do mammalian cells. To be sure, many biochemical and metabolic reactions are shared among different organisms but meaningful "prediction" of effects in the human cannot be made solely or primarily based on lower organisms.

During the last 20 years, the cosmetic industry has been pressurized not to carry out safety testing in animal's eyes. As a result , several programs have been launched looking for surrogate tests for ocular toxicity. Hopefully, the use of intact animals will be reduced with some of these procedures, particularly when one is working with the same cosmetic repeatedly. Unfortunately, no matter how well meaning or how dedicated research scientists or other individuals may be, there is no true absolute substitute even for an organ as relatively simple as the intact eye. When one tests for toxicity in the eye with a new drug candidate, one measures a spectrum of effects, beginning with the determination of whether or not the substance "burns" or "stings," which is manifest by the animal blinking or pulling away. In addition to burning or stinging, one determines whether or not the substance causes redness or tearing and whether there is any physical damage done to the surface of the eye or, via absorption, to the inner segment of the eye. If an animal's eye is anesthetized at the time the medication is administered, one cannot tell whether or not it burns or stings because the animal will make no response, since it cannot feel a burning or stinging sensation. If cells are removed from the eye, those that are capable of growing or surviving *in vitro* can only respond in a static way to substances that are put on them and the intact eye is not static. Note the continuous flow of tears and other lubricants, the continuous blinking that removes substances from the surface of the eyes, the exposure to oxygen, the ready availability of macrophages and other cells that digest foreign materials that penetrate the eye, the need for eye-to-brain-to-eye "communication," etc., and one sees the risk inherent in concluding that, by putting drug candidates on cells in culture, one will be able to accurately predict whether they will show any form of toxicity in the intact eye. The same comments pertain for all other organs of the body. To be sure, surrogate *in vitro* tests must be actively sought but, when new chemical entities are being studied, some evaluation in the intact animal eye will, in my opinion, still be required. In spite of existing hurdles, the *in vitro* surrogate test methods currently under study for the eye are well worth pursuing and hopefully will reduce the number of intact animals used for routine testing of some drug candidates, particularly when they are structurally related to compounds already known to be safe.

Three figures are included to further illustrate the complex communication system among cells in the body and within a given cell. Figures 16.1 to 16.3 show the tremendous complexity of metabolic reactions going on in the estimated tens of trillions of cells in the body at all times, while the heart recirculates ca 7200 L (1800 gallons) of blood throughout the animal on a daily basis (see Table 16.1). It has been suggested that one million cells die per second in the human. Although, as shown in these figures, a great deal is known about the illustrated pathways, it is a certainty that much more goes on in the live tissues in the body than we understand. Such complexity requires study of new drug candidates in whole animals to assess the total impact of a given compound on the intact body.

The use of computers is frequently highlighted by activists for generating useful information that will permit the scientist to bypass lower animals. As noted above, we know too little about all the reactions taking place in the intact animal body to be able to "calculate" our way around them. Computers are used wherever applicable in many areas of research. To be sure, if a company is working within a chemically related series of compounds and several have been tested for a given activity (be it efficacy or toxicity), one may well be able to test these compounds for primary and surrogate endpoints and computerize the data in

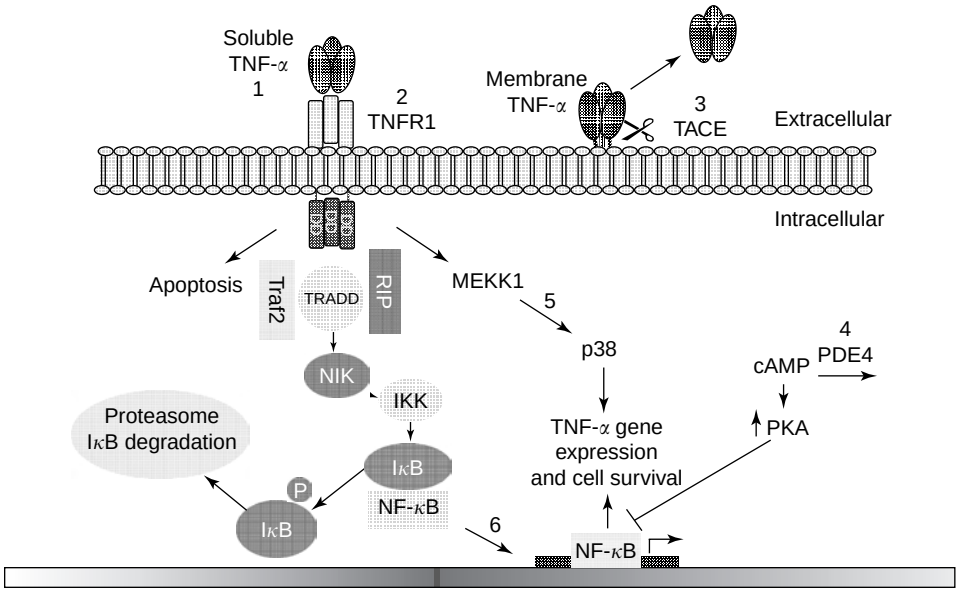


FIGURE 16.1
Signaling pathways. (From Bahjat, F.R., Characterization and Genetic Analysis of the Resistance of Nonobese Diabetic Mice to Tumor Necrosis Factor-Alpha Mediated Hepatocyte Apoptosis and Lethality, A dissertation presented to the Graduate School of the University of Florida in partial fulfillment of the requirement for the degree of doctor of philosophy, University of Florida, 2002. Reproduced with permission.)

an attempt to predict what effects the next compound in that chemical class will have on the animal. Such experiments are carried out regularly in pharmaceutical laboratories, often with considerable success. As valuable as such data are when working within a given congeneric chemical series, the studies may shed no light whatsoever on the reactions to be expected when a totally new chemical structure is introduced into the biological milieu. Fortunately, the highly developed computer programs in use today are leading to the design of unique, active molecules by techniques of gene mapping and proteomics and this science is very eagerly applied in pharmaceutical laboratories. Expectations are high that these investigations will yield new drug candidates with fewer side effects than has been possible in the past.

Another apparently simple, but yet complicated situation is the example of the tolerance to pain of medication given by injection. Some activists believe that all experiments in live animals should be carried out only in anesthetized subjects. If one uses anesthesized animals to test for tolerance after intramuscular or intravenous injection, one can determine whether or not there is any physical damage at the site of injection after the fact, but one cannot determine whether the injection itself is painful because the subject is anesthetized. Some substances, when injected into the muscle, are intensely painful and must either be abandoned, administered with a local anesthetic, or modified in some other way so as to permit their tolerance in human beings, especially babies and small children. Does anyone have the right to demand that all new injectable medications be given first to sentinent humans to determine their degree of any resulting pain? Surely the answer must be too obvious to discuss further.

The subject of the use of pets in animal research is often raised. The vast majority of work that involves dogs is carried out in animals that are bred for that purpose. If dogs are obtained from an animal shelter, it is most frequently for use in short-term experiments in which the animal is fully anesthetized, instrumented, and euthanized, over a period of

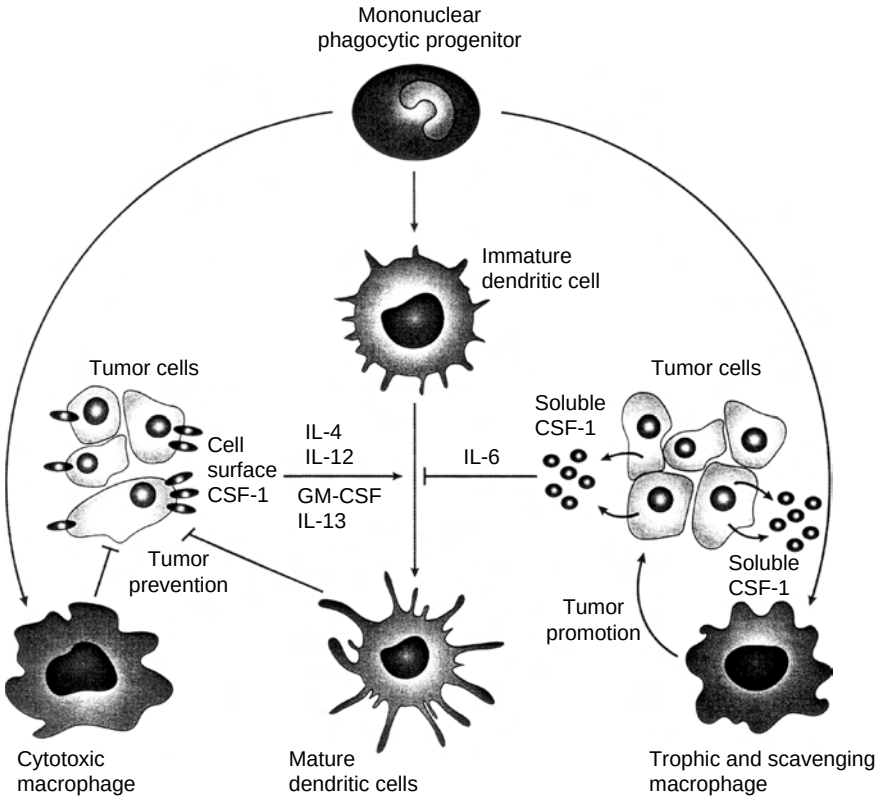


FIGURE 16.2
Tumor-educated macrophages. (From Pollard, J.W., Tumor-educated macrophages promote tumor progression and metastasis, *Nat. Rev. Cancer*, 4, 71–74, 2004. Copyright year 2004, Macmillan Magazines Ltd and the author. Reproduced with permission from Macmillan.)

hours. My estimate is that some 95% of dogs are purchased for research purposes from licensed breeders. Greater than 90% of all animals used in research are rodents. In 1989, the NIH noted that around 13 to 15 million animals were euthanized each year because they had no owners and facilities, and finances available to maintain them.[9] The so-called Pound Laws, which prohibit the sale of animals from public shelters for research purposes, actually increase the total number of animals ultimately sacrificed because replacements will be purchased from breeders and those in the shelters will, in the vast majority of cases, be euthanized. An excellent summary of the subject of animals in research has been published.[10]

No responsible human being can be indifferent to the suffering of sentinent animals under any circumstances. The activists have done animals and science a service by raising awareness of the need for proper controls on animal housing and use in all laboratories. They have done great disservice to the society, medical science, and animals (that also receive medications to treat their diseases) by the continued pressure to reduce or stop the use of animals in biomedical research. Certain of their demands may sound reasonable to individuals who are unfamiliar with the process of scientific discovery. For example, the demand to limit or eliminate "repetitious" experiments and not approve those that seem redundant to the inexperienced observer is based on totally fallacious thinking. Two scientists can do almost the same experiment but obtain results at different points in time, or at the same time in different laboratories, and reach different conclusions. For example,

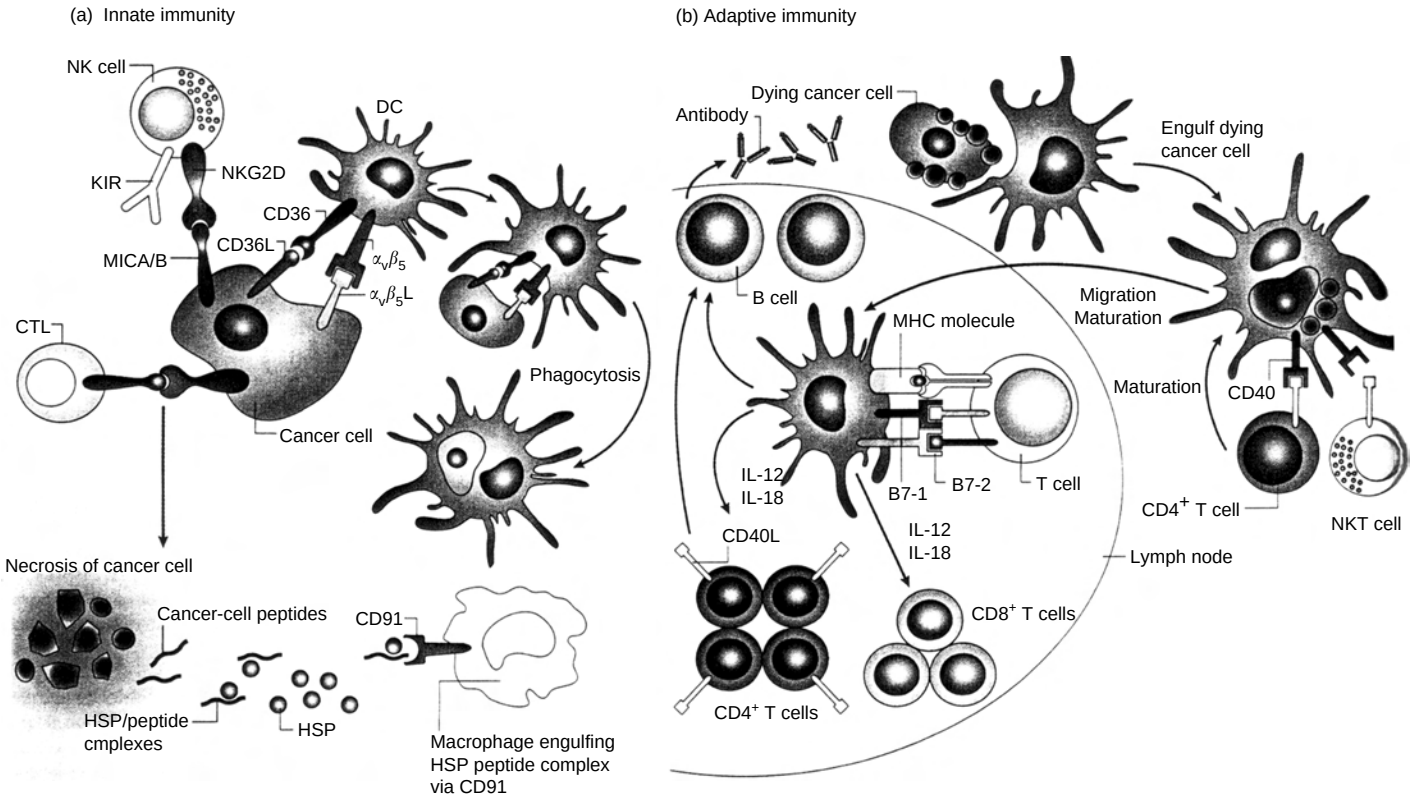


FIGURE 16.3
Cellular interactions. (From Dranoff, G., Cytokines in cancer pathogenesis and cancer therapy, *Nat. Rev. Cancer*, 4, 11–22, 2004. Copyright year 2004, Macmillan Magazines Ltd and the author. Reproduced with permission from Macmillan.)

suppose an effect occurs 36 h after a certain manipulation is carried out in the laboratory. A scientist who is observing the experiment at 36 h will see that change, whereas one who examines the subject at 24 or 48 h will not. When one reads the initial protocol upfront, the conclusion might well be that these are "repetitious" experiments and only one should be approved. Can any committee be so clairvoyant as to know when an effect will be seen in an experiment that has not yet been performed? Of course not! Furthermore, confirmation of laboratory findings is a *sine qua non* in science and simply cannot be abridged if we are to have confidence in the work reported.

If the sad day ever comes when animals are markedly reduced or prevented from being used in biomedical research for drug discovery and development, we will be witnessing a terrible day in the history of health progress in society and the flow of new drugs to treat the serious disease problems of today (AIDS, cancer, Alzheimer's disease, Parkinson's disease, hepatitis C, etc.) will drop to a trickle or dry up altogether. Society must not be so blinded by "feel good" comments and well-meaning but misguided sentiment for animals as to permit such a catastrophe to befall mankind.

References

1. Aziz, T., The contribution of animals to human health and wellbeing, Lecture to the British Parliamentary and Scientific Committee, December 13, 2004.
2. Jenkinson, N., Nandi, D., Miall, R.C., Stein, J.F., and Aziz, T.Z., Pedunculopontine nucleus stimulation improves akinesia in a Parkinsonian Monkey, *Neuroreport*, 15, 2621–2624, 2004.
3. Furth, J. and Kahn, M.C., Transmission of leukemia of mice with a single cell, *Am. J. Cancer*, 31, 276–282, 1937.
4. Skipper, H.E., Schabel, F., and Wilcox W.S., Experimental evaluation of potential anticancer agents on the criteria and kinetics associated with "curability" of experimental leukemias, *Cancer Chemother. Rep*., 35, 1964.
5. Holland, J.F., Historical perspectives, in *Neoplastic Diseases of the Blood*, 3rd ed., Wiernik, P.H., Canellos, G.P., Dutcher, J.P., and Kyle, R.A., Eds., Churchill Livingstone, Inc., New York, 1996, chap. 10.
6. Smith, C.G., Poutsiska, J.W., and Schreiber, E.C., Problems in predicting drug effects across species lines, *J. Int. Med. Res*., 1, 489–503, 1993.
7. Smith, C.G., Grady, J.E., and Northam, J.I., Relationship between cytoxicity in vitro and whole animal toxicity, *Cancer Chemother. Rep*., 30, 9–12, 1963.
8. Greaves, P., Williams, A., and Eve, M., First dose of potential new medicines to humans: how animals help, *Nat. Rev. Drug Discov*., 3, 226–236, 2004.
9. Goodwin, F., National Institutes of Health, 1989.
10. Vane, J.R., The discovery of drugs and animal research, in *Animal Experimentation and the Future of Medical Research*, Botting, J.H., Ed., Portland Press, London, 1992, chap. 6.

Appendix

For the reader to fully appreciate the working philosophy of the animal activist groups in the country, several quotes from their members are included here.

Quotes from: Animal Rights The Inhumane Crusade by Daniel T. Oliver, Merril Press, Bellevue, WA. Reproduced with the approval of the Capital Research Center, Washington, DC.

Michael W. Fox, vice president, the Humane Society of the United States:

> "The life of an ant and that of my child should be granted equal consideration." — *Inhuman Society* (1990 book)

> "We are not superior. There are no clear distinctions between us and animals." — *Washingtonian*, February 1990

Dan Matthews, celebrity recruiter. People for the Ethical Treatment of Animals
On the consequences of ending the use of animals in biomedical research:

> "Don't get the diseases in the first place, schmo." — *USA Today*, July 27, 1994

Ingrid Newkirk, president. People for the Ethical Treatment of Animals

> "I wish we all would get up and go into the labs and take the animals out or burn them down." — speech at the Animal Rights 1997 National Convention
> "Even if animal research produced a cure (for AIDS), we'd be against it." — *Vogue*, September 1989

> "In the end, I think it would be lovely if we stopped this whole notion of pets altogether." — *Newsday*, February 21, 1988

> "Animal liberationists do not separate out the human animal, so there is no rational basis for saying that a human being has special rights. A rat is a pig is a dog is a boy. They're all mammals." — *Washingtonian*, August 1986

> "Six million Jews died in concentration camps, but six billion broiler chickens will die this year in slaughter houses." — *Washington Post*, November 13, 1983

Alex Pocheco, chairman. People for the Ethical Treatment of Animals

> "Arson, property destruction, burglary and theft are acceptable crimes when used for the animal cause." — *Gazette Mail* (Charleston, VW). January 15, 1989

> "We feel that animals have the same rights as a retarded human child because they are equal mentally in terms of dependence on others." — *New York Times*. January 14, 1989

Peter Singer, author, *Animal Liberation*

> "There will surely be some nonhuman animals whose lives, by any standards, are more valuable than the lives of some humans." — Animal Liberation

17

Defining the Actual Research Approach to the New Drug Substance

Charles G. Smith

The approaches to discovering a new molecule discussed in this chapter will focus primarily on a molecule that is prepared by company scientists. Obviously, new drug candidates, either in early stages of development or nearing the market, can be available as the result of licensing activities, which will also be discussed in this chapter.

Taking an oversimplified approach to the question of exactly what type of molecule is desired and exactly how such a molecule might be designed, the following series of questions should be addressed:

1. Is the company's interest in "small" molecules (e.g., synthesizable by medicinal and organic chemists) or in macromolecules or products from natural sources?
2. If they are small molecules, should the objective be a new and unique chemical class, a "me-too," or both?
3. If it is a macromolecule, will the approach be a biotechnological one to produce known proteins that have not been available in the past in sufficient quantity to study or does the company want to produce analogs of the naturally occurring macromolecules?
4. In the same vein, does the company want to isolate a small molecule or macromolecule from fermentation sources, plants, or blood plasma to develop as a drug?

Obviously, the resources to be put in place will vary with the answers to the above questions. If specific macromolecules are selected as the product target, the major initial effort is the preparation of the desired macromolecule or molecules. In this instance, there will not be a large number of compounds from which to select. In the case of isolation of active leads from fermentation or plant sources, considerable efforts are devoted early on to "screening" (e.g., rapid, broadly based testing in a variety of biological systems) of the raw materials. Pure, biologically active compounds are isolated from the crude, naturally occurring sources, based on specific biological or biochemical assays in which the extracts were originally found to be active or, when available, using a surrogate chemical or physicochemical analytical method. In the case of a specific macromolecule to be prepared by biotechnological approaches, the ultimate pharmacologic activity is usually already known and at least one major application will clearly have been identified. In the case of materials isolated from natural sources other than blood, be they plants, marine sources, fermentation liquors, etc., a major and exciting potential advantage that resides in such molecules is that their structures may well be unique. In my experience, novel chemical structures (regardless of their

origin) have the greatest potential to open totally new therapeutic areas, once they have been proven to be safe and efficacious. At the same time, the more unique the chemical structure, the less one can predict its efficacy, metabolism, and toxicity. Accordingly, unique undesirable effects are not infrequently found in such molecules and one must balance the risk of failure against the potential major gain inherent in such programs.

A word about medicinal products from human blood plasma is in order. In the past, products such as antibodies and albumin have not been associated with the transmission of disease to patients, whereas antihemophilic products were known to carry a risk of transmitting the hepatitis virus. Since the recognition in the early 1980s of the AIDS virus (HIV) as an extremely dangerous agent that is transmitted sexually and in whole blood or in certain blood products, a lot of research has been done in the blood fractionation field: (1) test the HIV status of the donated blood so that it will not be pooled with uninfected blood prior to fractionation and (2) modify the processes used to prepare plasma fractions so as to kill or greatly attenuate any AIDS virus particles therein. The "AIDS scare" has also intensified biotechnology research aimed at producing various plasma proteins (e.g., factor VIII) *in vitro* that cannot be contaminated with HIV, since they were not derived from human blood. It is very unlikely that any new companies will enter the field of plasma fractionation as a source of new drugs in the future.

With respect to the acquisition of synthetic chemicals, there are several quite independent sources thereof: (1) compounds that one can purchase from the catalogs of various chemical companies; (2) compounds that one can obtain from universities or from other companies, usually on license; (3) compounds that are available from the NIH or other government laboratories; (4) new chemical entities (NCEs) that are designed by the chemists in the laboratory of the founding company (the so-called combinatorial chemistry approach will be discussed in detail in Chapter 12); and (5) "me-too" efforts that are directed at capturing some fraction of a large existing market by synthesizing a patentable analog or derivative of a known active agent. The sourcing of compounds that already exist from third parties (other than catalog purchases, which are usually done to test a concept and not done with the hope of finding a drug lead *per se*) is rather straightforward. The two primary approaches for acquisition of compounds are either through the research division or through the company's licensing group. All companies large enough to have a serious interest in licensing (either in-licensing or out-licensing) should have at least one individual in research and development (R&D) who is assigned the responsibility of providing liaison with the business unit in which the primary licensing function resides.

For compounds that may be available on a research basis within universities or companies, most frequently an appropriate person in R&D will contact a research colleague in academia or in another company asking for a sample of the desired compound for laboratory testing. Small samples for such purposes are frequently exchanged among R&D groups with no commitment whatsoever to licensing as a result thereof. For compounds that are already marketed or for those outside research compounds that are shown to be of interest by the research division and are nominated as in-licensing candidates, it is most usually the licensing executive at the company who contacts his or her counterpart at the organization that owns the compound to determine whether or not licensing discussions are possible. Once in-licensing targets have been established within a company, the licensing executive and appropriate representative from R&D should lay out a plan for contacting various companies or academic centers that might have compounds, of the types desired, available for consideration. Appropriate contacts are then made at the companies or academic institutions and meetings are arranged to discuss compound availability in detail. Obviously, one need only research catalogs (or an appropriate computerized sourcing list) to be able to purchase most commercially available substances.

A chemical that can be purchased from a catalog cannot, needless to say, be patented as composition of matter by the company that purchases it. The primary value of such compounds is generally to test a structure–activity hypothesis, from which chemists and biochemists then move forward to design more specifically active molecules. Reading the literature carefully and attending scientific conferences will keep scientists apprised of some compounds that are available from academia, but direct contact with the business development offices of major universities will, not infrequently, alert interested parties to available substances that may not be widely known in the literature.

In my experience, a good route to obtaining compounds for screening, in cases where a given company has a good test system in hand but lacks sources of chemicals, is to "door knock" at companies that made thousands of compounds over the years and have them sitting in inventory. If one tries diligently enough, one can usually find a source of interesting heterocyclic chemicals sitting on a shelf that have not been tested in a particular area of biological screening in which the licensee is interested. Such approaches to compound acquisition cannot sustain a large company in the long term, however, and are primarily valid during "startup" periods. Once the research team hits its stride, a continuous flow of compounds, the majority of which are prepared in-house, should be the objective of the chemistry group. High volume approaches such as combinatorial chemistry will be detailed elsewhere in this book. Small- and medium-sized companies undoubtedly will continue to bring new chemicals in from third parties. The very broad coverage via the Internet is widely used for searching.

The design of "me-too" compounds may not be very exciting to the bench scientists but, when successful, this approach has provided, in the past, products to the company on an intermediate term basis. Scientists must realize that, as their division grows and consumes more and more dollars for its operation (which, by its very nature, requires long time frames for success), the sales and marketing people in the company must continue to generate increasing revenues to keep the stock analysts and investors happy and to pay for research. When requested by the marketing group, a certain percentage of "me-too" work should be undertaken by research in close cooperation with marketing, since some of the products that result therefrom often generate very nice revenues for the company. This approach definitely enhances the perception of the value of the research division in the minds of the marketing, sales, finance, and corporate groups and, indeed, serves as a source of satisfaction to the scientists who see the fruits of their labors successfully treating patients and generating revenue for the company many years earlier than they could ever have expected from an unrelated NCE. At the same time as one willfully pursues "me-toos," the research management must convince the corporate management that this approach to new products must not be the major objective of a sophisticated research team, as it is not possible to keep good scientists motivated when the vast majority of their work is of a "me-too" nature. Selecting the "me-too" target area is usually not difficult as this process is almost totally market-driven.

In light of the major "merger & acquisitions" activity (M&A) over the past decade, large companies usually cannot generate sufficient revenue from "me-too" drugs to sustain the necessary growth rate of the company. As a result, most large pharma groups need new and unique products in major markets. Also, M&A activities bring together the chemical collections of the companies involved, which results in NCEs to all parties involved. Coupling this availability with combinatorial chemistry actually can markedly expand the drug candidate horizon of each company in the merger.

Last, but certainly not least in the milieu of finding a new drug candidate, is in-house synthesis of unique chemical entities. Companies that introduce truly unique NCEs into major markets: (1) often enjoy the satisfaction of being able to ameliorate or cure diseases that, before the advent of their NCE, could not be adequately treated; (2) enjoy a major

surge in prestige among their peers, physicians, patients, and shareholders as a result of this accomplishment; and (3) make considerably more profit for the company from the patented NCE than ever would have been possible with a "me-too" drug for use in the same disease.

Both the time-honored approaches of medicinal chemists (e.g., variation of molecules, combining of different molecular types into one new molecule, combinatorial approaches, etc.) and the highly sophisticated science of computerized drug design are used to prepare NCEs. The computational chemistry approach follows two major pathways: (1) it attempts to design a molecule or synthesize a bioactive segment or fragment of a known active molecule directly and (2) approaches to determine the three-dimensional (3D) molecular site of biological interest (e.g., receptor site, active enzyme, etc.) to which one wishes to bind a ligand or an inhibitor and to design the small molecule based on the known structure of the 3D site. The new areas of genomics, proteomics, and antisense research, discussed in other chapters of this book, are expected to yield exciting new drugs. All new molecules that are prepared by the chemist are evaluated in the screening test pertinent to the disease area in which the chemists are assigned to work. When supplies are adequate, all NCEs should be screened as broadly as available material will allow, to permit the serendipitous discovery of new biological activities, which has played an important role in drug discovery in the past. Unless a company has a very large chemical synthesis group, high-volume screening tests (e.g., able to handle several thousands of compounds) will be able to outstrip the supply of compounds prepared in-house. Screening capacity that is available over and above that needed to accommodate those compounds that are synthesized in-house for the specific project should be saturated from existing libraries of chemicals within the company that have not yet been evaluated in the particular test under consideration or by candidates that are brought in from the outside, as discussed above. All initial biological activities are confirmed by retesting, followed by dose–response determinations in both *in vitro* and in whole-animal test systems (discussed in more detail in Chapter 4). Chemists are well aware of the need for patent action and, in all chemical approaches to prepare new compounds, adequate worldwide patent coverage is of critical importance.

Over the last decade, a very interesting approach to the discovery of new molecules with desirable biological activities has been introduced. In such programs, a totally random series of chemical reactions ("combinatorial libraries") gives use to hundreds or thousands of compounds that are screened in desired biological test systems. Unbelievable numbers of new compounds can be synthesized using such approaches, which will be discussed in detail in Chapter 12. The ability of this technology to increase the rate of new drug discovery is yet to be proven. In addition to synthetic chemicals as a source of medicinal agents, the newly expanded fields of biotechnology research (for proteins) and oligonucleotide synthesis (for antisense molecules) also provide exciting new molecules. These subjects are discussed in more detail in Chapters 11 and 12.

Compounds from natural sources can also form the basis for NCEs. Natural sources include fermentation liquors (also known as "beers"), plant extracts, and extracts of marine organisms. In the fermentation approach, scientists isolate microorganisms of interest to them in large numbers and carry out sophisticated, rapid, high-volume screening in which thousands of the xenobiotics (biologically produced molecules that show biological activity in a particular test system) produced by the organisms are screened, in impure mixtures, for the desired biological activity. When an activity is found, the organism (called a "culture") is grown in flasks and fermentation tanks to prepare enough of the fermentation liquor to provide a source of crude material from which pure compounds must eventually be separated. In such an effort, methods have to be put into place to identify *known* xenobiotics early on, so as not to "reinvent the wheel" on a regular basis. In some percentage of the fermentation beers investigated, new molecules that

possess interesting biological activities will be identified. Once a pure compound is isolated, the chemists determine the structure of the molecule so that studies aimed at synthesizing the parent substance and preparing synthetic derivatives and analogs thereof can be carried out. As noted in a previous publication, the structures of compounds prepared from fermentation sources are so unique that, at times, they can be expected to "defy synthesis and defeat rationale." This statement applied equally well to compounds isolated from plant sources. A company interested in the fermentation approach must be prepared to invest considerable sums of money in fermentation equipment, as it is necessary to ultimately produce thousands of liters of beer to recover adequate material for full biological, toxicological, and clinical testing and ultimate marketing.

For fermentation screening, companies collect soil samples from all over the world in the hope of including widely differing types of organisms in their primary screens. Some laboratories attempt to increase uniqueness in the organisms they isolate by superimposing an "isolation pressure" on the soil samples that they process. For example, a soil may be collected from under an asphalt roadbed and be plated out on a growth medium in which the only carbon source is a hydrocarbon. By definition, only microorganisms that can oxidize a hydrocarbon (and presumably soil under an asphalt road has been enriched by years of biological selection so that it contains significant numbers of organisms that can do so) will grow out on the plate. Such "selection pressure" should permit the microbiologist to examine organisms that would not normally be isolated when a standard, rich laboratory medium is used for the primary growth phase of the organisms in the soil sample, since the organisms in largest numbers that are capable of growing rapidly in nutritionally rich medium will simply outgrow those that have unique nutritional requirements. In the example given, the hydrocarbon nutrient will allow only those organisms that are adapted to metabolizing hydrocarbons to be isolated. Once such specimens are grown out, the cultures are replated onto richer media in which the organism can be expected to produce its xenobiotics. Another interesting source of unique organisms is soil that has been heavily irradiated in the past and in which the microbial flora has been reestablished (e.g., atomic bomb testing sites and Brookhaven National Laboratory sites). Obviously, an infinite number of variations around the theme of bioselection pressures can, and should, be employed.

Plants provide another important source of NCEs for new drug discovery research. Indeed, the pharmaceutical industry may be looked at as having its origin in plants, plant extracts, and chemicals from plants. In spite of this fact, the yield of new medicinal agents from plant sources in recent decades left a great deal to be desired outside the antitumor field. There are many reasons for this, including (1) the fact that a large number of valuable drug leads can be produced by synthetic and other chemical approaches, such that the biological test systems downstream are saturated without the need to collect plants, transport them, extract them, etc.; (2) the availability of the desired plant often cannot be assured on second or third collections and a given plant specimen may actually appear to be quite different at different times of collection; (3) yields of the desired material may vary widely with different batches of the plant or at different times of the year; and (4) collection, shipment, and exportation and importation across borders may provide significant barriers that are not experienced by synthetic chemists or companies that have fermentation facilities in place, operating under their own close control. In spite of these difficulties, plants remain a rich source of unique compounds with biological activity, and certain interesting and creative new approaches to the exploration of plant sources are currently being pursued. Marine sources of new molecules share many of the problems that have been experienced in seeking new chemical leads from plants. "Marine sources" refer to seaweeds that grow underwater and to actual fish, crustaceans, etc. *per se* and not solely the result of microorganisms that might be isolated from aquatic sources. The

reason for this statement is that microorganisms, once isolated, become part of the "pool" of microorganisms that may be obtained from a wide variety of sources, and the isolated organism moves directly into the "fermentation" screen.

Isolating new potential drug candidates from human blood plasma is a highly specialized business, as noted above. Only those companies that are invested in facilities and personnel to collect and fractionate human blood can participate in this activity. In recent years, the fear of contamination of the plasma with the HIV (AIDS) or hepatitis viruses and the tremendous advances in molecular biology and biotechnology (e.g., genetic engineering) that permit the synthesis of body proteins outside of the animal body have combined to discourage new entries into the plasma fractionation field. While it is unlikely that companies currently in this business will totally abandon the approach, it seems very unlikely that any new ventures will be undertaken therein.

Having considered sourcing molecules for evaluation as potential drug candidates, the drug discovery process now moves on to screening and secondary biological evaluation, discussed in the next chapter.

18

Pharmacokinetics — Pharmacodynamics in New Drug Development

Sarfaraz K. Niazi

CONTENTS

Part 1: Good Pharmacokinetic Practice to Streamline New Drug Development

18.1 Story of a Drug Molecule

I am a drug molecule; I can cure and I can kill. My birth is a tedious and burdensome process — financially and chronologically (my gestation time now hovers around 12 years) — so much so that only a few well-to-do companies can afford to bear the pain. On average, I cost about a billion dollars to become a reality, regardless of what I am capable of doing or what I should not be doing (how toxic I am) and regardless of whether I am a run-of-the-mill analgesic or a novel protease inhibitor intended to annihilate the acquired immunodeficiency syndrome (AIDS) virus. U.S. companies spend over $50 billion per year to look for more prospects like me. Such monumental costs are a result of two things: regulations and legal exposure. Because of the mishaps caused by drugs such as thalidomide, regulations enacted since 1962 now require that I go through comprehensive testing — a testing so thorough that when I am given to millions, the effects and toxicity ratios should not change. (You see, all of us are toxic because we are mostly foreign to the body; it is when the good we do surpasses what we should not be doing that we become an acceptable risk.) Drug companies working on me are aware of all this, but what forces them to pay heed to the regulations is legal liability; one death — one new side effect — can wipe me and the company working on me out of the market if I am not properly documented — with a trail of diligent studies to show that this outcome was out of question.

18.1.1 The Discovery Journey

Discovering me begins in many different ways: Most drug companies have scientists sitting in cluttered suites thinking about what I should look like — the so-called drug-modeling approach — and then they go out and construct me out of simpler chemical entities. The combinatorial chemists can work on mass-scale approaches for discovering me, studying various permutations and combinations of chemical components. Armed with robotic controls and gigabytes of churning read-only memory, they can now synthesize a large number of my look-alikes in a much shorter time period. It is no longer unusual for a company to synthesize 100,000 possible hits in a year and screen them for probable activity. I can also be brought to life by various fermentation processes or from bacterial sources, particularly now, with the advent of genetic engineering. Biotechnology has opened a new vista for companies looking for my variants, and when they find one, bacteria such as *Escherichia coli* make an ideal turf for me to grow into a crop. The newest, and the most promising, search for me is found in proteins — particularly small proteins. Most recurring diseases, the scientists

believe, are a result of genetic defects; proteins can avert the effects of many genetic disorders. As a result, a new field of study has opened up; small-protein science has its own experts — both analytical chemists and drug designers — all looking for me. Having found my probable structure, companies begin to understand how and why I work as I do. Although direct testing on animals is still the primary method of studying how I act, several newer techniques — such as inducing cell cultures and *in vitro* drug–receptor interactions — can also help them understand me.

More traditionally, scientists discover me by reverting to the age-old proverbial indications: the herbal source. Given the long, successful history of finding new brethren of mine — such as aspirin, quinine, reserpine, and taxol — in the roots and leaves of trees, this approach seems promising. Drug companies must secure a patent on my life to justify the millions (and now close to a billion) they must spend to raise me into a full-grown product on the shelf; without that intellectual property right protection, I will never see the light of the day.

18.1.2 The Regulatory Maze

The U.S. Food and Drug Administration (FDA) — the granddad of all regulatory authorities — makes it easy for everyone to know what to do when working with me; however, doing what FDA wants is a tough and dollar-laden track. First, I must be fully characterized — physically and chemically, with such attestations as infrared spectra, mass spectra, crystallographic structure, solubility, stability, and a host of other tests — which takes about 6 to 9 months to complete. Having fully characterized me, companies repeat pharmacological screening (the first screening is often done on a less pure form) to find out if I have retained the activity I am supposed to have. Then begins the cycle of toxicology testing: teratogenicity and carcinogenicity are the main tests besides the famous LD_{50} (how much of me it takes to kill half of the animals).

Having satisfied the regulators with the toxicity profile in animals (three species) and armed with a detailed pharmacokinetic (PK) profile, which includes detailed pharmacokinetic–pharmacodynamic (PK/PD) and physiologically based pharmacokinetic (PB/PK) or marker-based studies among many other variants, researchers begin their planning to put me into humans; however to do so, they need a dosage form. The dosage-form development, therefore, begins simultaneously with my toxicity-potential testing to cut the time shorter, but it is an expensive exercise.

At this stage, the companies file an Investigational New Drug (IND) application with FDA, requesting permission to test me, for the first time, in humans. The FDA grants this approval on the basis of relative weight of my efficacy potential vs. my observed toxicity in animals. FDA's desire to modernize itself requires that it approve the IND within 30 days, putting the onus of timeline responsibility on the regulatory authorities. Unfortunately for my sponsors, they can expect to hear from FDA on the 25th day after filing, and that phone call is not likely to be good news, by the way. To allow testing in humans requires a critical judgment for two reasons: there is no guarantee that my observed toxicity in animals will be replicated in humans because there are no perfect models to simulate toxicity in humans — it is always a risky matter — and FDA judges the potential of efficacy. If I am a treatment for cancer or AIDS, then FDA will tolerate greater toxicity because the diseases themselves can be fatal; however, if I am a simple analgesic, I would better be made really safe to get FDA to approve me. (Some people say that if aspirin were discovered today, then FDA would reject it for its use as an analgesic.)

18.1.3 Meeting Lots of People: The Clinical Journey

With an approved IND, the drug company can place me in humans at lower doses, mainly to observe any overt symptoms of toxicity and to evaluate my PK characteristics, and the study may use 10 to 20 healthy subjects. These data are then resubmitted to FDA (Phase I). After my Phase I is completed, drug testing begins for my clinical efficacy, with trials at different doses and given to different populations under strict supervision. A successful study may require a few thousand patients and the involvement of several study centers across the country (or perhaps across the world) to collect extensive data. It is not unusual to have 100,000 patients take me before the data are submitted to the FDA for drug approval — the so-called marketing approval.

Marketing approval, when received, is hailed as a major victory for the drug company, which has invested heavily. The marketing decision is quickly reflected in the drug company's stock market price per share, either upward or downward, depending on FDA's decision. With marketing approval, I emerge as a new drug — beating the one in 5000 odds of a proven, efficacious compound resulting from all the "good leads." Some say, the odds are worse than winning on a roulette wheel in Las Vegas; but people do, sometimes, come out of a Las Vegas casino smiling.

This presents the quick snapshot of how I came to life. The real story — of how I work, how I cure a disease, and how I can kill — fills volumes of binders. Also lying buried somewhere in those volumes are the scientific rivalries, intellectual property fights, failed departmental projections, and nervous notes from bankers.

18.1.4 Breaking Out of My Shell: The Dosage-Form Story

Just why market approval is such a victory has a lot to do with the environment I must cross. So, let us go back and look at my form again. My journey from identification as a new chemical entity to becoming a product on the shelf requires a detailed study of how I work, not only in the body but also in dosage form. Investigations on me began with knowing which end of me is straight (crystal-wise) — all in an attempt to determine whether I like water (hydrophilic) or not (hydrophobic) because if I do not like water, I will not dissolve, and I would, therefore, have difficulty crossing the many biological barriers I must go through to reach the spot where I am needed, the site of action.

I am also studied to see whether I have the prowess to withstand the hostile environment of the human body; the environment I must survive can dissolve an ingot of iron, and I must traverse it from the stomach to intestine, from which I will, most likely, be absorbed. If the scientists cannot find a dosage form that will make the environment friendly toward my survival, I will die prematurely; there are scores of very potent drugs that will never see the light of the store shelf because they are unstable and a proper dosage form could not be designed for them. So, the discovery of a new activity is only half the battle; getting that activity to work in the body is, perhaps, equally challenging and often more disappointing.

Having placed me in a dosage form, my journey begins when the body gets to work on me, before I can work on the body. If I am a solid dosage form (tablets, capsules, and the like), which is most often the case, then I must be able to come out of the tight environs of my form and into the environs of the body fluids. It is at this stage that a lot of problems in bioavailability are encountered, because if I do not come out of my form quickly and completely, I will pass down the intestinal column to areas in which, even if I do come out, I may not be absorbed into the body. Dosage-form design, therefore, becomes a crucial part of my usefulness. This is why studies on determining drug activity must involve designing dosage forms simultaneously.

18.1.5 Traps, Mines, and Barriers: Getting to the Action Site

After coming out of my dosage form, I face the formidable task of avoiding many hijackers en route to the site of activity: Mucin and binders of other kinds may trap me just long enough so that I get pushed down the canal to the area from which I cannot be absorbed. I may also be deactivated chemically at this stage. And this is the point at which I must demonstrate that I like both water and oil because I must pass through two barriers; if I prefer one, I may not get through the other barrier. This journey is like crossing mine barriers.

If all worked just right, then I have, through a miracle, reached the blood circulatory system with 10 to 70% of my original quantity. The blood inevitably goes through the organ of my destruction: the liver. This is the first survival test I must pass. Liver is there to keep me out of the system, and so, it immediately begins to knock me down; the key is to lose as few of my soldiers here as possible. In fact, this battle with the liver is often beneficial and intended: My breakdown product may be the active form of the drug (in which case, I began as a prodrug).

Having crossed the liver, I face another challenge: how to sail through the blood vessels. Often, I have the option of moving through the blood vessels by tagging along with plasma protein molecules. This, however, can be a double-edged sword. If I can tag along, I may develop a love affair with the proteins and never make it to where I am needed. Binding to proteins helps me stay in the body longer, however, and the bathing of tissues with plasma allows me to enter strange valleys through many unique barriers: blood, brain, cerebrospinal fluid, tissues, and the like.

So a few of us, finally, get to reach the place where we are needed: the site of action. The irony is that few scientists know where that site really is. They often adduce a plausible theory about where I am, but the full facts of my journey are rarely known because I can have overlapping activities across various sites. I may work on the brain when my main job is to kill bacteria in the tissue; yet, my activity in the brain might be helping to boost immunity, nevertheless. It is rarely important how many of my molecules reach the site of action; what is important is how long I stay there. And that is the job of a pharmacokineticist: to evaluate and modify my design to give me a longer lease on life.

18.1.6 Bidding Adieu: Making My Exit

My purpose is to help, but then like all other good armies, I must also leave when the job is over. The human body is a true marvel of nature. It has many weapon systems ready to destroy me. Often, I leave the body through the kidneys; other times, I recirculate and end up in the intestines or in other body tissues. But eventually, I am knocked out, although some of us are truly hard and stick around for years, especially when we develop an affection for proteins. My exit from the body is measured using many mathematical techniques, and many parameters are assigned to me on the basis of statistical principles that I know very little about, such as the Bayesian theory, deconvolution modeling, dose–response correlations, and the like: These are the topics of discussion in this book.

This is a brief history of my sojourn through the human body. To fully understand my potential and my usefulness, scientists must invoke two disciplines: the physicochemical and the clinical. It is the marriage of these two that makes it possible to make the best use of me. So, biopharmaceutics and clinical PKs must be studied in unison. Those working on me must know how the population handles me, the physiologic basis of my disposition, and how the time course of my life cycle is related to my activity and toxicity. Somehow, the regulatory agencies are always very curious about me.

18.2 Introduction and Background

The regulatory submissions have, during the past 25 years, become more demanding as a result of the greater emphasis placed on the safety of new drug molecules. A significant component of the safety profile of a drug and its recommended dosing regimen is based on the PK characteristics of the drug. With our improving understanding of human physiology, the use of biological markers (biomarkers), and the availability of better computing tools, highly sophisticated mathematical models can be invoked to support the safety profile of new drug molecules. One goal of the clinical pharmacology section of a regulatory filing is to describe and predict the relationship between drug–dose and drug–effect. Pharmacokinetic profiles of a new molecule describe the factors affecting the dose–active site concentration process, and PD describes the activity site concentration–effect process.

The term *pharmacokinetics* refers to the way a drug is handled by the body. It measures parameters such as the area under the curve (AUC), the concentration at maximum (C_{max}), and other parameters calculated from those measured parameters, such as clearance, half-life, and volume of distribution, which reflect the absorption (A), distribution (D), and elimination (E) of a drug from the body. A drug can be eliminated either by metabolism (M), as one or more active or inactive metabolites, or by excretion of the unchanged drug through various routes, notably the kidneys. The overall set of processes — often referred to as absorption–distribution–metabolism–excretion (ADME) — ultimately controls the systemic exposure of the body to a drug and its metabolites after drug administration. This relationship is often extended to the release or liberation (L) of drug molecules from the dosage form: LADME. This systemic exposure, reflected in plasma drugs and metabolite concentrations, is generally used to relate the drug dose to both its beneficial and adverse effects. All drugs show inter- and intra-individual variances in PK measures and parameters, compounding the mathematical difficulties in making projections about the time course of drug concentration and activity in the body.

18.3 The Exposure–Response Relationship

The exposure–response relationship is critical for regulatory review because it determines the safety and effectiveness of drugs. A drug can be determined to be safe and effective only when the relationship between the beneficial and adverse effects *vis-à-vis* an exposure to the body is known. Drug molecules can range from the well-tolerated drugs with little dose-related toxicity, in which the PK properties from a single dose to multiple doses are linear, to toxic molecules, in which clinical use must be determined on the basis of a nonlinearity in their PK properties. Sometimes with the toxic drugs, the doses can be titrated for tolerability, but it is always important to develop information on population exposure–response relationships for favorable and unfavorable effects and to adjust the exposure for various subsets of the population. Pharmacokinetic characterization in pediatric cases is of particular importance to FDA.

The term *exposure* refers to the amount of drug entering the body, and various measures of observed or integrated drug concentrations in the plasma and other biological fluids (such as C_{max}, C_{min}, C_{ss}, and AUC). *Response* refers to a direct measure of the pharmacologic effect of the drug and includes many endpoints or biomarkers: clinically remote biomarkers (such as receptor occupancy), a presumed mechanistic effect (such as ACE inhibition), a potential or accepted surrogate (such as effects on blood pressure,

lipids, or cardiac output), or the full range of short-term or long-term clinical effects related to either efficacy or safety.

The choice of dose, dosage form, type of drug release, the dosing interval, and how the drug is monitored depends on the magnitude of an effect and the time course. Exposure–response and PK data further clarify this relationship *vis-à-vis* intrinsic and extrinsic patient factors, such as genetic phenotypes and age effects, for instance. Additionally, exposure–response data can be used as well-controlled clinical studies to support effectiveness; to emphasize the efficacy, especially where surrogates are used; and to support dosing in populations other than those studied. It is important to know that, in several situations the relationship between dose and plasma concentration is rather poor, primarily because of nonlinear PK results or inter-individual variations; in such situations, the systemic exposure studies offer the most valuable supporting tool for validating the efficacy profile of a new drug molecule. For example, it might be reassuring to observe that even patients with increased plasma concentrations (metabolic outliers or patients on other drugs in a study, for instance) do not have increased toxicity in general or with respect to a particular concern (such as QT prolongation on electrocardiograms). However, determining blood concentration remains indispensable, particularly when both the drug and its metabolites are active, where different exposure measurements (such as C_{max} and AUC) provide different relationships between exposure and efficacy or safety, when the number of fixed doses in the dose–response studies is limited, and where responses are highly variable and the data are intended for an exploration of the underlying causes of the response variability.

Exposure–response data, using short-term biomarkers or surrogate endpoints, can sometimes make further exposure–response studies from clinical endpoints unnecessary. For example, if it can be shown that the short-term effect does not increase beyond a particular dose or concentration, there may be no reason to explore higher doses or concentrations in the clinical trials. Similarly, short-term exposure–response studies with biomarkers might be used to evaluate early (i.e., first dose) responses seen in clinical trials.

Exposure–response information for a new target population can sometimes be used to eliminate the need for clinical studies by showing comparable concentration–response relationships using short-term clinical or PD endpoint. The same holds true for dose and dosing-interval changes, formulation changes, and the like. In some cases, if there is a change in the mix of parent and active metabolites from one population (e.g., pediatric vs. adult), dosage form (because of changes in a drug-input rate, for instance), or route of administration, additional exposure–response data with short-term endpoints can support the use in the new population, the new product, or the new route without further clinical trials.

In new formulations, exposure–response data are used to isolate unintentional PK differences (e.g., in the release profile and, thus possibly, the metabolic profile through the saturable metabolic system). With these data on hand, a company can use *in vitro* or *in vivo* bioavailability studies as sufficient proof of equivalence, particularly between the formulations used in establishing the initial efficacy proof and the formulation intended for marketing. In biological drugs, changes in the manufacturing process often lead to subtle, unintentional changes in the product, resulting in altered PKs. When a change in the product can be determined as not having any pharmacologic effects (i.e., no effect on unwanted immunogenicity), exposure–response information may allow appropriate use of the new product. The comparability protocols suggested by FDA, and recently by European authorities, heavily emphasize on these studies for biological drugs.

Exposure–response data, however, are not likely to obviate the need for clinical data when formulation or manufacturing changes result in altered PKs, unless the relationships between measured responses and relevant clinical outcomes are well understood. The sponsors of studies need to make a financial evaluation of these studies as well; often,

while working with regulatory authorities, the choice of studies submitted depends on what would be the least-expensive alternate. In my experience, many such cost-saving measures end up actually costing more; the key is to provide the most robust data supported by multiple studies to account for most of the plausible variables.

In measuring exposure–response relationships, the choice of active moiety is important; ideally, it should include the active parent and active metabolites, particularly in those instances where the metabolism is route-dependent (first-pass metabolism). In such instances, hepatic and renal impairments become important considerations. There are many important considerations in selecting one or more active moieties to measure in plasma and in choosing specific measures of systemic exposure. When drugs are optically active and administered as the racemate, the differences in the PK and PD properties of the enantiomers should be elucidated to ascertain whether using one of the pure enantiomers as the final drug product might be valuable.[1]

Complex drug substances can include drugs derived from animal or plant materials and drugs derived from traditional fermentation processes (yeasts, molds, bacterium, or other microorganisms). For some of these drug substances, identification of individual active moieties or ingredients is difficult or impossible. In such a circumstance, measurement of only one or more of the major, active moieties can be used as a "marker of exposure" in understanding exposure–response relationships and can even be used to identify the magnitude of contribution from individual, active moieties. A special case of protein binding is the development of antibodies into a drug. Antibodies can alter the PK of a drug and can also affect PK/PD relationships by neutralizing the activity of the drug or preventing its access to the active site. This is of major concern in the development of biogeneric drugs because many biological drugs are coming off patent. The fact that antibodies produced by these molecules can result in sensitizing the body toward its indigenous production of the product (e.g., insulin or erythropoietin) and that there is no good way to measure, this has kept the regulatory authorities from accepting biogeneric applications. In March 2006, generic human growth harmone was approved.

Pharmacokinetic concentration–time curves for a drug and its metabolites are used to identify primary exposure metrics such as AUC, C_{max}, or C_{min}, which are not time-dependent unlike the sequential measurements of concentration over time. A peak plasma concentration of a drug is often associated with a PD response, especially with an adverse event. There can be large inter-individual variability in the time-to-peak concentration, and closely spaced sampling times are often critical to determining the peak plasma concentration accurately in individual patients because of differences in demographics, disease states, and food effects, if any. All these elements are clearly spelled out in the protocols written to conduct these studies.

During chronic therapy, collection of multiple plasma samples over a dosing interval is often not practical. As a substitute, a trough plasma sample can be collected just before administration of the next dose at scheduled study visits. Trough concentrations are often proportional to AUC because they do not reflect drug-absorption processes as peak concentrations do in most cases. For many of the drugs that act slowly relative to the rates of their ADME, trough concentration and AUC can often be equally well correlated with drug effects.

Although collection of serial plasma samples and measurements of the corresponding response is the most desirable method, in some instances this process may not be ethical or possible, such as in pediatric or geriatric populations. An alternative method is to obtain plasma samples at randomly selected times during the study, or at prespecified but different times, to measure drug concentration and, in some cases, response. Although this type of sampling will not allow the usual PK data analysis for making precise estimates of individual PK parameters, the use of a specialized technique — population PK analysis

combined with the Bayesian estimation method — can be used to approximate population and individual PK parameters. This will allow evaluation of exposure variables that can be more readily correlated to response than the few plasma concentrations studied.

The efficacy and safety of a drug can be characterized using a variety of measurements or response endpoints, such as clinical outcomes (clinical benefit or toxicity), effects on a well-established surrogate (change in blood pressure or QT interval), and effects on a more remote biomarker (change in ACE inhibition or bradykinin levels). In many cases, multiple response endpoints are desirable especially when less-persuasive clinical endpoints (biomarkers, surrogates) are used in choosing doses for the larger and more difficult clinical endpoint trials. Greater problems arise in studies conducted across many sites or where multiple laboratories are used to measure the outcome.

The regulatory authorities place great emphasis on exposure–response studies. Poorly designed studies are likely to confound essential properties of drug molecules; as a result, the protocols must be written with preciseness and clarity to prevent misinterpretations.

18.4 Biomarkers

Biomarkers can provide great predictive value in early drug development if they reflect the mechanism of action for the drug, even if they do not become surrogate endpoints. Data from genomic and proteomics that differentiate healthy from disease states lead to biomarker discovery and identification. Multiple genes control complex diseases via hosts of gene products in biometabolic pathways and cell and organ signal transduction. Pilot exploratory studies are required to identify pivotal biomarkers for predictive clinical assessment of disease progression and the effect of drug intervention.

Most biomarkers are endogenous macromolecules, which can be measured in biological fluids. Many biomarkers exist in heterogeneous forms with varying activity and immunoreactivity, posing challenges for bioanalysis. Reliable and selective assays can be validated under something like a good laboratory practice (GLP) environment for quantitative methods. Whereas the need for consistent reference standards and quality control monitoring during sample analysis for biomarker assays is similar to that for drug molecules, many biomarkers have special requirements for sample collection that demand well-coordinated team management. Bioanalytical methods should be validated to meet study objectives at various drug development stages and should perform adequately at quantifying biochemical responses specific to the target disease progression and drug intervention. Protocol design for producing sufficient data for PK/PD modeling would be more complex than that for PK alone.

Knowledge of the mechanism for discovery and of the preclinical studies is helpful in planning clinical study designs in cascade, sequential, crossover, or replicate modes. The appropriate combination of biomarker identification and selection, bioanalytical methods development and validation for drugs and biomarkers, and mechanism-based PK/PD models for fitting data and predicting future clinical endpoints and outcomes provides powerful insight and guidance into effective, efficient, rational drug development toward safe and efficacious medicines for individual patients. The use of biomarkers and surrogate endpoints of efficacy, hazard, and exposure in preclinical studies has evolved rapidly in validating projections to drug response on the basis of PK studies.

Biomarkers are a variety of physiologic, pathologic, or anatomic measurements that relate to normal or pathological biologic processes.[2,3] These include characteristics — such

as physical signs, blood analytes, physiological measurements, and the like — that are objectively measured and validated as indicators of normal biological processes, pathogenic processes, or pharmacological responses to the use of a drug. The use of biomarkers presupposes that the cause or the susceptibility or the disease progression is related to a biomarker measurement and that this correlation is applicable at the clinical level (from pharmacologic effect to clinical response). Biomarkers differ in how close they come to serving the intended purpose:

- Some biomarkers are valid surrogates for a clinical benefit (e.g., blood pressure, cholesterol, and viral load).
- Some reflect a pathologic process and are, at least, candidate surrogates (e.g., brain appearance in Alzheimer's disease, brain infarct size, and various radiographic or isotopic function tests).
- Some reflect a drug action but have an uncertain relationship with the clinical outcome, such as the inhibition of platelet aggregation dependent on adenosine diphosphate or ACE inhibition.
- Some biomarkers are more remote from the clinical benefit endpoint (e.g., the degree of binding to a receptor or inhibition of an agonist).

Regulatory authorities do not accept a biomarker as an acceptable, surrogate endpoint for the determination of a new drug's efficacy unless it has been empirically shown to function as a valid indicator of clinical benefit (i.e., it is a valid surrogate). This relationship cannot be built on a theoretical justification for leaving biomarkers to support but not replace surrogate endpoints; this is because many biomarkers will never undergo the rigorous statistical evaluation that would establish their value as a surrogate endpoint. The time course of changes in biomarkers is often different from changes in clinical endpoints as they more closely approximate the time course of plasma drug concentrations; therefore, the arguments presented against using plasma concentration as a measure of efficacy apply to biomarkers as well. However, biomarkers do offer an alternative in establishing the exposure–response relationships for dose range in clinical trials intended to establish efficacy and provide insight into potential adverse effects. Biomarkers can also be useful during the drug discovery and development stages, where they can help link preclinical and early clinical exposure–response relationships and better establish dose ranges for clinical testing.

Often, drug responses are difficult to obtain (such as that of a bacterial cure from an antibiotic), are difficult to measure quantitatively (such as the mood elevation produced by an antidepressant), have variable effect on longevity (for instance, in the survival time from a cancer therapy), or are unethical to measure (a necropsy score for a safety evaluation, for instance). Therefore, the effect of ultimate interest in a PK/PD trial may be replaced by a surrogate endpoint. The clinical validity or relevance of a surrogate is determined by its statistical association and mechanistic links with a clinical outcome (surrogate accuracy). In addition, the surrogate should have desirable metrological properties, namely, reproducibility (of measurement), continuity (for a graded quantitative measurement), objectivity, specificity, and linearity.

Surrogate endpoints, a subset of biomarkers, are laboratory measurements or physical signs used in therapeutic trials as a substitute for clinical endpoints expected to predict the effect of the therapy.[2] A fully validated, surrogate endpoint predicts the clinically meaningful endpoint of a therapy consistently.[3]

FDA is able to rely on less well-established surrogates for accelerated approval of drugs that provide meaningful benefits over existing therapies for serious or life-threatening illnesses, such as AIDS, rare carcinomas, or breaking infectious diseases. In these cases, the

surrogates are likely to predict clinical benefit on the basis of epidemiologic, therapeutic, pathophysiologic, or other scientific evidence. Surrogates generally would not be used to evaluate clinical relevance or side effects when these effects are unrelated to the surrogate.[2] Examples in which prospective and retrospective trials in human medicine have demonstrated statistical correlations between the surrogate markers and either clinical success or prevention of resistance include indices mechanistically related to clinical outcome because they are all constructed using the minimum inhibitory concentration (MIC) values. For ACE inhibitors (such as benazepril and enalapril) that help prevent heart failure, PK/PD relationships have been investigated using plasma and tissue ACE inhibition. On the basis of these relationships, dosage regimens have been designed.[4] As is with the case of AUC and MIC or C_{max} and MIC, ACE inhibition is only a surrogate endpoint — survival time and quality of life are the ultimate goals, which are, nevertheless, difficult to monitor in ACE-inhibition therapy.

18.5 Adding Modeling to New Drug Development

There is broad recognition within the pharmaceutical industry that the drug development process, especially the clinical component of it, needs considerable improvement to keep pace with changes in the requirements and needs of the health care environment.[5] Modeling and simulation are mathematically validated techniques that have long been used extensively in disciplines other than the pharmaceutical industry (for instance, in the automobile, aerospace, and computer design industries) to develop products more efficiently. Both modeling and simulation rely on the use of (mathematical and statistical) models that are simplified descriptions of the complex systems under investigation. Other types of models that are becoming increasingly important (and which are discussed more fully in Part 2) are population models, stochastic simulations, and ongoing efforts to integrate models for disease progression and patient behavior (such as compliance).

For both modeling and simulation, a guidance on "good pharmacokinetic practice" — analogous to GLP, good clinical practices, and good manufacturing practices — is currently evolving through the joint efforts of academia, regulatory agencies, and industry (as described elsewhere in the book). These guidelines will be a great impetus for model-based analysis and simulation, contributing to the streamlining of the pharmaceutical drug development processes.[6]

Part 2: Pharmacokinetic/Pharmacodynamic Modeling in New Drug Development

18.6 Introduction

Drug development identifies the right molecule for the right target using a dosage regimen that yields the optimal therapeutic outcome. Safety information and adequate and well-controlled clinical studies that establish a drug's effectiveness are the basis for the approval of new drugs. The risk inherent in the administration of any new drug entity depends on

the nature and extent of exposure to the body; this information is pivotal in determining the safety and effectiveness of drugs. A drug is determined to be safe and effective only when the relationship of beneficial and adverse effects with a defined exposure is known. There have been few instances, and they are now getting rarer as more potent molecules are designed, in which a drug is well-tolerated with little dose-related toxicity by all patients and, therefore, single-dose studies can be used to represent the full nature of exposure, taking into account PK and idiosyncratic parameters. More often, the drugs are very toxic, and their clinical use can only be based on weighing the favorable and unfavorable effects at a particular dose or doses. Whereas in some instances, the doses can be titrated to effect or to tolerability, in most cases, it remains important to develop information on population exposure–response relationships for favorable and unfavorable effects and information on how, and whether, exposure can be adjusted for various subsets of the population such as pediatric, elderly, or genetically compromised population.

The critical exposure–response data generally derived from the preclinical and clinical studies provide a basis for integrated model-based analysis and simulation.[7,8] Simulation allows us to predict an expected relationship between exposure and response in situations where real data are absent or difficult to obtain. Models often do not establish a causal relationship or provide an explanation for the mechanism of action for a drug, and as a result, they may not be used as a basis for approval of a new drug; they can, however, help in analyzing data from well-controlled clinical trials. To predict an exposure–response relationship, well-controlled clinical studies that investigate several fixed doses or that measure systemic exposure levels are required by which it is possible to analyze, using scientifically reasonable causal models, and validate hypotheses about the effects of alternative doses and dosage regimens that are not actually tested. This information is useful in optimizing dosage regimen and individualizing treatment in specific patient subsets for which limited data are provided.

In a clinical response, there are two major sources of variation: the PK and the PD of the drug. Applying computer-based methods to the population PK/PD approach helps to separate these sources of variability.[9] Mixed-effect models analyzing population PK/PD data explain variation between subjects (or groups of subjects). These models are capable of handling pooled (often sparse) data while allowing for fixed or random effects. Improving the power of the estimation process for the PK and PD parameters improves the suggested optimal-dosing regimens.

Because all the reasonable dosing regimens cannot be studied using the current standard of a 48-week controlled study of efficacy and safety, the goals of finding an optimized dosing regimen are best achieved by integrating PK (describing the relationship between dose and concentration vs. time) and PD parameters (describing the relationship between concentration and effect vs. time), often though a link model (bridging the PK and PD models) and a statistical model describing the intra- and inter-individual variability. When a PK/PD model is employed, both the time course and the variability (in the effect vs. time relationship) are predicted for different dosage-regimen scenarios. However, mechanistic PK/PD models can be relied upon for extrapolation, i.e., prediction vs. description. A prior determination of a safe and effective dosage regimen for use in pivotal clinical trials always proves beneficial, both in terms of cost and risk in clinical trials.[10] Several elements are crucial for successful PK/PD modeling and simulation for efficient and effective, rational drug development, including:

- Mechanism-based biomarker selection and correlation to clinical endpoints.
- Quantification of drug and metabolites in biological fluids, under GLP.
- GLP-like validation of the biomarker method and measurements.
- Mechanism-based PK/PD modeling and validation. This involves the four distinct steps of building PK model, building PD model, linking PK and PD models, and simulation of treatment regimens or trials for useful prediction.

18.6.1 Assumptions in Pharmacokinetic/Pharmacodynamic Modeling

In the process of PK/PD modeling, it is important to describe, prospectively, the objectives of the modeling, the study design, and the available PK and PD data. The assumptions of the model can be related to dose–response, PK, PD, or one or more of the assumptions listed in Table 18.1.

The assumptions can be based on previous data or on the results of any available current analysis. What constitutes an appropriate model depends on the mechanism of the drug's action, the assumptions made, and the intended use of the model in decision-making. If the assumptions do not lead to a mechanistic model, an empirical model can be selected, in which case, validating the model's predictability becomes especially important. (Note that nonmechanistic models do not get good reviews from the FDA.) The model-selection process comprises a series of trial-and-error steps, in which different model structures or newly added or dropped components to an existing model can be assessed by visual inspection and can be tested using one of several objective criteria. New assumptions can be added when emerging data justifies it.

18.6.2 Model Validation

Model validation is a process that involves establishing the predictive power of a model during the study design as well as in the data analysis stages. The predictive power is estimated through simulation that considers distributions of PK, PD, and study-design variables. A robust study design will provide accurate and precise model-parameter estimations that are insensitive to model assumptions.

During the analysis stage of a study, models can be validated on the basis of internal or external data. A common method for estimating predictability is to split the data set into two parts: build the model on the basis of one set of data and test the predictability of the resulting model on the second set of data. The predictability is especially important when the model is intended to provide supportive evidence for primary efficacy studies, to address safety issues, to support new doses and dosing regimens in new target populations or subpopulations defined by intrinsic and extrinsic factors, or when there is a change in dosage form or route of administration.

TABLE 18.1

Assumptions in Pharmacokinetic/Pharmacodynamic Modeling

A PK/PD Model Can Be Related to Dose–Response, to PK, to PD, or to One or More of the Following Assumptions	
The mechanism of the drug actions for efficacy and adverse effects	Presence or absence of active metabolites and their contribution to clinical effects
Development of tolerance or absence of tolerance	Immediate or cumulative clinical effects
Disease state progression	Drug-induced inhibition or induction of PK processes
Circadian variations in basal conditions	Response in a placebo group
Absence or presence of an effect compartment	Influential covariates
The PK model of absorption and disposition and the parameters to be estimated	The PD model of effect and the parameters to be estimated
Inclusion or exclusion of specific patient data	Distributions of intra- and inter-individual variability in parameters
Distribution of PK and PD measures and parameters	

18.7 Pharmacokinetic Modeling Studies

18.7.1 Compartment Pharmacokinetic Modeling

Pharmacokinetics is the study of the movement of drug molecules in the body, requiring appropriate differential calculus equations to study various rates and processes. The rate of elimination of a drug is described as being dependent on, or proportional to, the amount of drug remaining to be eliminated, a process that obeys first-order kinetics. The rate of elimination can, therefore, be described as

$$\frac{dX}{dt} = -kX \tag{18.1}$$

where k is a mere proportionality constant or a rate constant and X is the amount remaining to be eliminated (and therefore X_0 is the initial amount or the dose administered). Integration allows converting Equation 18.1 to

$$X = X_0 e^{k_{el}t} \tag{18.2}$$

Because the amount X is proportional to the concentration, a similar equation describes the time-decay profile of the drug concentration instead of the amount:

$$Cp^t = Cp^0 e^{-k_{el}t} \tag{18.3}$$

This simple, first-order relationship allows a linear association between the log (more appropriately, the natural logarithm) of concentration and time. It is noteworthy that this concentration is the "effective" concentration and not necessarily the measured concentration. *Effective* refers to a thermodynamic activity rather than the physical concentration. Drugs decay in proportion to the concentration of "free" drug molecules, and whatever is bound to proteins may not be available for disposition. This extrapolation becomes more complex when we take into account other factors that might alter the "activity" (in a thermodynamic sense) of the drug in a biological fluid. For example, structuring of water inside protoplasm imparts lipophilic characteristics, which create significant differences in the available concentration gradients. This is a primary reason as to why it is not always possible to correlate measured concentrations with pharmacologic responses because the level of drug at the site of action or at the receptors depends highly on the thermodynamic activity of the drug, which is difficult to assess.

The relationship between the amount of drug and its concentration is classically represented by the following equation, which functions as if there were a physical space (called distribution volume) throughout, which the drug distributes evenly:

$$V = \frac{\text{amount of drug in the body}}{\text{concentration measured in plasma}} \tag{18.4}$$

This relationship is an oversimplification of the distribution characteristics of drug molecules in the body and can provide results in volumes often much larger than the body weight. For example, if a drug were selectively stored in different parts of the body, such

as digoxin or diazepam are, the apparent distribution volumes, using Equation 18.4, would be several multiples of the body's weight. Because the distribution of a drug is a time-dependent process, even within the same "compartment," it is suggested that this parameter be treated as a time-dependent variable[11] (as we will see later); treating a "bolus" dose as a short-term infusion improves the results of the deconvolution of integrated equations. This assumption allows a more accurate physical representation of the PK models because an "instantaneous" intravenous injection is treated as a very short duration, zero-order, input function. As we shall see, this consideration is more important as we integrate PD models where the action and effect of the drug is delayed for several reasons, including the input and distribution variables.

The area under the plasma-concentration time curve, the AUC, is a useful parameter in defining the overall body exposure to a drug; this parameter integrates the concentration-over-time function:

$$\text{AUC} = \int_{t=0}^{t=\infty} Cp^t \, \mathrm{d}t \tag{18.5}$$

Because the time function of drug concentration is dependent on the rate at which the drug is cleared from the hypothetical "volume," the AUC function is dependent on total body clearance, CL:

$$V = \frac{\text{dose}}{\text{AUC} \cdot k_{\text{el}}} \tag{18.6}$$

Clearance (CL) is a product of volume, V, and the elimination-rate constant, k_{el}, such as when the drug is removed from the urine or metabolized or removed from the sampled compartment by another means. This description of clearance often confuses students of PK. Clearance is an inherent phenomenon, in which distribution volumes are high and rate constants are small to compensate for the distribution. Both the volume and the rate constant are derived phenomenon and do not determine clearance. Note that total body clearance is a composite of all pathways that clear or remove the drug from the sampled compartment or the compartment from which the drug is cleared; this is based on the mathematical relationship between the observed elimination-rate constant and its components: each of the pathways is involved in the turnover of the drug within the body. Using the parameters described above, it is possible to "simulate" a sampled compartment (of fluid) concentration as a function of time in a single- or multiple-dose application using simple, iterative programs. Numerous computer programs are now available, which are drug- and model-specific, and which allow simulations of steady-state blood levels that depend on various body functions and body characteristics that affect the clearance of the drug. Mixed models, involving a zero-order infusion, a bolus, or other similar combinations, can be made to estimate blood concentrations under different circumstances related to drug administration.

When drugs are received by routes other than intravenous injection, input is not "instantaneous" or a short-order zero, and the function must often be represented as a mixed-order, primarily a first-order, process, which must then be taken into account in simulating drug concentration. Drug clearance, however, is not always a constant parameter, especially when an organ such as the liver is involved in the removal of the drug from the body:

$$\text{Organ clearance} = \frac{Q(C_{\text{a}} - C_{\text{v}})}{C_{\text{a}}} = QE \tag{18.7}$$

where Q is the blood-flow rate to the organ, C_a is the concentration of the drug in the blood when entering the organ (in the arterial blood), and C_v is the concentration of drug in the blood when leaving the organ (in the venous blood). The term E is the steady-state extraction ratio. High E values mean high clearance by the liver and, thus, extensive metabolism. The liver blood-flow rate is a physiological parameter that can be altered in disease states. The extraction ratio depends not only on the function of liver but also on the nature of the drug. Both the hepatic clearance and the extraction ratio are empirical parameters and depend on the total hepatic blood flow, the unbound fraction of the drug, and the intrinsic clearance rate. Intrinsic clearance is differentiated from total clearance; the former is the ability to transform when other factors are not present. In other words, intrinsic clearance is the property of a body organ that clears the drug such as liver or kidney; for example, the maximum clearance in kidneys cannot exceed the total blood rate to the kidneys and the hepatic clearance cannot exceed the total blood flow to the liver. The actual clearance of a drug from the body depends on the intrinsic clearance as well as its concentration in the fluid that is being cleared; a lower concentration resulting from distribution to body tissues will reduce the total clearance but will have no effect on the intrinsic clearance:

$$\mathrm{CL} = Q\frac{\mathrm{fu}\,\mathrm{CL}_{\mathrm{int}}}{Q+(\mathrm{fu}\,\mathrm{CL}_{\mathrm{int}})} = \frac{Q\,\mathrm{CL}_{\mathrm{int}}^{\mathrm{total}}}{Q+\mathrm{CL}_{\mathrm{int}}^{\mathrm{total}}} \tag{18.8}$$

This makes the extraction ratio

$$E = \frac{\mathrm{fu}\,\mathrm{CL}_{\mathrm{int}}}{Q+(\mathrm{fu}\,\mathrm{CL}_{\mathrm{int}})} \tag{18.9}$$

High-clearance drugs are those for which there is no saturation of the reaction that converts the drug, and therefore, the clearance rate approaches the blood-flow rate. For capacity-limited drugs, flow rate is irrelevant, and clearance is a simple product of the unbound fraction and the intrinsic clearance.

The traditional method of PK data analysis uses a two-stage approach: estimation of PK parameters through nonlinear regression using an individual's extensive concentration–time data and using these data parameters as input data for the second-stage calculation of descriptive summary statistics on the sample. These statistics typically include the mean parameter estimates, and the variance and the covariance of the individual parameter estimates. Analysis of dependencies between parameters and covariates using classical, statistical approaches (linear stepwise regression, covariance analysis, cluster analysis) can be included in the second stage. The two-stage approach yields adequate estimates of population characteristics. Mean estimates of parameters are usually unbiased, but the random effects (variance and covariance) are likely to be overestimated in all realistic situations. Refinements have been proposed (such as the global, two-stage approach) to improve the traditional approach through bias correction for the random effects of covariance and differential weighting of individual data according to the data's quality and quantity.

18.7.2 Physiologically Based Pharmacokinetic Studies

Physiologically based PK studies take a different perspective in modeling drug disposition in human body — a mechanistic physiologic distribution model. This approach had been in use in other disciplines long before the compartment kinetic modeling was applied for

TABLE 18.2
Perfusion Rates of Human Body Organs

Organ	% Body Weight	Perfusion Rate (L/min/kg)
Fat	20	0.056
Muscles	48	0.03
Brain	2.0	0.56
Skin	3.7	0.12

studying drugs. In 1937, the mathematical basis for physiological PK modeling was established by Teorell,[12,13] but the solution to the equations was too difficult to obtain before the invention of the digital computer. An automatic solution of a physiologically realistic, mathematical description of the uptake, distribution, and clearance of a chemical agent was proposed by Bischoff in the early 1960s.[14] At that time, computation limitations forced several simplifications to the models, including the assumption that the distribution of the drug between tissues and blood is instantaneously at equilibrium, which led to physiological models with blood flow–limited delivery of chemicals to tissues. The inhalation PK models using instantaneous distribution are well-known. Physiological PK studies progressed no further until the early 1970s, when the physiological parameters of human organ system became better known and digital computers became more-widely available. Today, physiological PK modeling is critical to understanding the behavior of a drug at the site of action.

Exposure modeling studies are often based on the physiologic functions that determine uptake, distribution, and elimination of drugs from the body. This approach was pioneered using anesthetics in which physical distribution determines both the onset and termination of action.[15] Similar results have been reported for other compounds such as D_2O and ethanol,[16] propranolol,[17] and inulin and protein-bound antibiotics.[18] The modeling is based on a quantitative description of the distribution process using standardized organ weights and blood-flow rates (Table 18.2). A simpler model assumes no solute binding and a tissue/plasma equilibrium coefficient of 10 for all tissues, except for muscles where this value is 3.62 and for fats where this value is 2.42 as used for propranolol.[19] Also in this simple model, there is no first-pass effect and kidney excretion is the only mechanism of drug removal from the body; thus, the input function is equal to systemic availability. In more complex models, tissue binding and other factors that produce nonequilibrium of the tissue/plasma ratios are introduced. The simple model, when used to determine bolus-response function, is well-described by a simple two-exponential function; in the more complicated models, three exponents generally provide good fit, and often, going to higher exponents does not improve the predictability. More important is the timing of the first data point obtained in the bolus-response function. This should, ideally, be obtained at or before the end of the constant infusion. (Note that better estimates are obtained from infusion studies than from single-bolus doses because there is always an inevitable delay in the dispersion of drug in the bolus dosing, but the model assumes no delay.) When a deconvolution method is used (see below), the robustness of analysis depends on the accuracy of venous-concentration data because the response function $r(t)$ is established from these data; therefore, any errors in this function reduces the reliability of the analysis, particularly when a later time sample, such as 10 min, is used as the first data point.

18.7.3 Bioequivalence and Systemic Exposure Models

Screening drug molecules for suitability for use in humans is often subjected to certain basic toxicity or workability solutions to reduce the cost of screening. The human body

must be able to remove the drug in a reasonable time. Drug clearance is an intrinsic parameter; however, body clearance (extent of drug removal) is dependent on cardiac output and the overall extraction ratio:

$$\text{Body clearance}(\text{plasma, blood}) = \dot{Q} \times \text{ER} \tag{18.10}$$

The extraction ratio (ER) ranges from 0 to 1, and the cardiac output is proportional to body size:

$$\dot{Q}(\text{mL/kg/min}) = 180\,\text{BW}(\text{kg})^{-0.19} \tag{18.11}$$

Cross-species comparisons can be made for crude estimates and, generally, for drugs that have clearance of less than 4 mL/min/kg would be evaluated only if there are special reasons that the mechanisms of actions need to be evaluated.

In addition to the removal potential of a drug, the entry potential is also a good screening parameter; for drugs that are poorly bioavailable, further development should proceed only if proper modification to the molecular structure or to the drug delivery system is made to provide a reasonable possibility of entry. When evaluating bioavailability, it is important to first establish a PK basis because of the large variation in bioavailability as a result of the differences in population PK. Population models are most appropriate for this type of evaluation. Obviously, the consideration of bioequivalence in establishing compliance of generic products is important, and the guidelines for these measurements are defined in the *United States Pharmacopoeia* and other guidelines provided by the FDA. It should be noted that the purpose of these studies is to compare the systemic exposure of the body to the drug molecules; this requires measurement of both the extent of absorption and the rate of absorption. Traditionally, parameters such as AUC, T_{max}, and C_{max} are studied using specified statistical models. For drugs given orally, these studies cannot be substituted with PD studies, which may be required for some drugs in which the plasma or sample tissue concentration is not available.

18.7.4 Deconvolution Techniques

The bolus-response function $r(t)$ is generally described using a multiexponential function:

$$r(t) = \sum_{i=1}^{\rho} a_i \left(e^{-t/T_i}\right) \tag{18.12}$$

The optimized values for a_i and T_i are determined by using mathematical approach without any significance attached to it for physiologic reasons.[20–22] Generally, the re-sorting required to use a three-exponential term takes the estimates out of the population parameters or global minimum.

The three parameters in γ-distribution are chosen by minimizing the error function:

$$\text{Error function} = \sum_i \frac{\left|(y_{\text{gam}})_i - (y_{\text{dat}})_i\right|}{(y_{\text{dat}})_i + \text{noise}} \tag{18.13}$$

where $(y_{dat})_i$ is the sum of overall data points for the experimental venous concentration and $(y_{gam})_i$ is the venous concentration-determined y convoluting the γ-distribution input

using a polyexponential equation as described above for $r(t)$. The *noise* factor in Equation 18.13 determines the weighting of each data point used. When there is no error of *noise*, then the error is simplified for each point. When the error is large, the term $(y_{dat})_i$ drops out in the denominator and the error is proportional to the numerator of the error term. Because the γ-distribution function is a highly nonlinear process, it is important to use a global annealing procedure such as that used in PKQuest (Minneapolis, MN; www.pkquest.com, requiring Maple software, Maplesoft, Ontario, Canada, www.maplesoft.com),[23] and then follow it with nonlinear minimization. The venous concentration is fitted by using interpolation, meaning that it goes through each data point or uses a smoothing cubic-spline function, and then, performing the deconvolution. The B-spline function defines the number and position of "breakpoints" and the order of the spline function. Highly sophisticated models have been used for this purpose.[24]

The course of systemic exposure to a drug is studied by comparing intravenous (IV) administration studies using deconvolution approach, in which $r(t)$ is the systemic concentration produced from IV administration (also called bolus function) and $I(t)$ is the systemic input rate (in units such as g/min) from the nonintravenous (non-IV) route:

$$c(t) = \int_0^t r(t-\tau)[I(\tau)]\,\mathrm{d}\tau \tag{18.14}$$

If there is no first-pass effect involved, then $I(\tau)$ is equal to the rate of intestinal absorption upon administration of equal doses (in IV and non-IV forms). In first-pass metabolism, $I(\tau)$ is the systemic availability of the drug upon oral (or sublingual, rectal, buccal, etc.) dosing. The function $r(\tau)$ is obtained by fitting the data upon IV administration to a variety of exponential equations and selecting the best fit through residual mean error of fit. The duration of infusion can be instantaneous (a few seconds for bolus input) but, more realistically, is usually a few minutes. Whereas it is desirable to obtain the sample as early as possible, sampling earlier than 2 min after injection is not advised, so as to allow time for venous mixing. Long-term IV infusions are also used to obtain the $r(\tau)$ function. Mathematical solutions of the deconvolution are easily obtained by using such validated software as PKQuest requiring Maple[23] software. Several methods are used for deconvolution; γ-distribution input is a parametric-fitting technique. Whereas polyexponential-fitting techniques are widely used, better fits are obtained by using a parametric approach for simulating $I(\tau)$, where A is the amount of drug reaching the circulation, Γ the γ-function, a the γ-number that ranges from 1 to 6, and b has inverse time units:

$$I(t) = \frac{(Ab)^a t^{a-1} \mathrm{e}^{-bt}}{\Gamma(a)} \tag{18.15}$$

This approach offers a superior simulation, particularly in situations where there is a delay in the input function such as in intestinal absorption and gastric emptying variations.[25] The three parameters (A, a, and b) given above are estimated by global (also called *simulated annealing*) and local (also called *Powell*) nonlinear optimization.[26] The fitting of data using γ-deconvolution method smoothes data noise, and with no user-adjustable parameters, the bias is removed. If it is not possible to define the input using a single γ-distribution, then other deconvolution approaches, such as analytical, spline, or uniform approaches, which remove the "roughness" of the input rate are used; the choice of parameters is additionally improved by experimental Akaike criterion[27] and the "generalized cross-validation."[28]

The analytical deconvolution involves approximation of $C(t)$ by an interpolating or smoothed spline function and the devonvolution.[26,29] The analytical deconvolution method is most commonly used for the advantage of being fast and also where data are exact, excellent results are obtained; however, the robustness of this approach depends on the value chosen for the smoothing parameter, which is poorly estimated even when standard deviation is available (very rare). Where there is noisy data, it adds more error in analytical deconvolution compared to spline and uniform methods. Also, analytical deconvolution does not allow use of negative values for input. In spline function input consideration, the input $I(t)$ is parametrized using a general B-spline function and then obtaining deconvolution by a constrained regression.[30,31] In using uniform input, $I(t)$ is estimated on dense uniform sequence of time points and then using stochastic regularization procedure for deconvolution.[24,32]

18.7.5 Population Pharmacokinetics

The main purpose of generating PK data is to simulate various disease conditions, physiologic variables, and patient characteristics in arriving at effective dosing regimen. This exercise is important because it is not always possible to collect data in all eventual situations or to obviate the need to collect extensive data that may not be possible. Planned population PK studies (see below) to test the design factors prove very useful. A simulation scheme requires repetitive simulations, followed by analysis of data sets, to account for the effect of sampling variability on parameter estimates. A permutation of study designs can be simulated to determine the most useful design, yielding the best information.

There are generally two types of protocols used for simulations: the add-on and the standalone. Both protocols contain clear statements of the population-analysis objectives and the proposed sampling design and data collection procedures. The specific PK parameters to be investigated are identified in advance. If the population PK study is added on to a clinical trial (an add-on study), as is done frequently, then the PK protocol should not compromise the primary objectives of the clinical study. Investigators should be made aware of the value of including a population PK study in a clinical trial.

The population model defines at least two levels of hierarchy. At the first level, PK observations in an individual (such as concentrations of drug species in biological fluids) are viewed as component of the individual probability model. In this model of individual observations, the mean distribution is represented by a fitted model such as a biexponential model where mean is given by a PK model. At the second level, the individual parameters are regarded as random variables and the probability distribution of these (often the mean and the variance, i.e., the inter-subject variance) is modeled as a function of individual-specific covariates. These models — their parameter values and the use of study designs and data analysis methods designed to elucidate population PK models and their parameter values — are what is meant by *population PKs*.

Population PK analysis is now pivotal during drug development for understanding the quantitative relationships among the drug-dosing profile, the patient characteristics, and the drug disposition. It provides a better understanding of the variability within a target population. Nonlinear, mixed-effects modeling proves very useful in dose-ranging studies (such as, so-called titration or effect-controlled designs), in which there is a reasonable expectation that inter-subject kinetic variations may require changing dosing regimens for some subgroups within the target population, such as when the population is heterogeneous or when the therapeutic window is narrow.

Pharmacokinetic characterization through estimation of PK parameters returns greater value if the parameters of disposition and their statistical distribution in larger

populations can be ascertained (as shown above in the two-stage modeling practice). Population PK is the study of the sources and correlates of variability in drug concentrations among individuals who are the target patient population receiving clinically relevant doses of a drug of interest. Patients even within a narrow demographic group are different in how their body acts on drug molecules (the PK part of the equation); this includes drug excretion, biotransformation, and distribution. Obviously, disease state and use of other medications have a significant impact on these disposition profiles and alter the dose–concentration response. Variations become larger when we compare age, sex, genetic and even environmental factors. For example, steady-state concentrations of drugs that are mostly eliminated by the kidney are usually higher in patients suffering from renal failure. For drugs with uncomplicated mechanisms of excretion, the kidney function would be a good measure to project half-lives of drugs in any population — the goal of population PK studies. However, a difference in the plasma concentration alone cannot be a predictor of drug response or drug toxicity; there must exist a relationship between concentration (or another surrogate) and the response to allow determining the window of variation allowed. Population PK seeks to identify the measurable pathophysiologic factors that alter the dose–concentration relationship and the extent of those changes, so that, if such changes are associated with clinically significant shifts in the therapeutic index, dosage is appropriately modified. These projections can be made (when subjected to appropriate statistical models) even when the amount of data is limited. Additionally, the analysis of data from a variety of unbalanced designs as well as from studies that are normally excluded because they do not lend themselves to the usual forms of PK analysis (such as concentration data obtained from pediatric and elderly patients or data obtained during the evaluation of the relationships between dose or concentration and efficacy or safety) can also be used in a population PK modeling.

The magnitude of the unexplained or random variability in PK evaluation is important because the efficacy and safety of a drug can decrease as unexplainable variability increases. In addition to inter-individual variability, the degree to which steady-state drug concentrations in individuals typically vary about their long-term average is also important. Concentrations may vary as a result of inexplicable, day-to-day or week-to-week kinetic variability or may be the result of errors in concentration measurements. Estimates of this kind of variability (residual intra-subject, inter-occasion variability) are important for therapeutic drug monitoring. Knowledge of the relationship among concentration, response, and physiology is essential to design the dosing strategies for rational therapeutics that may not necessarily require therapeutic drug monitoring.

Whereas the variability in PD parameters is likely to be higher than what the PK would predict, given the highly variable inter- and intra-individual effects, the cross-species comparisons are more likely to vary because of PK effects. For example, the binding of drug can be different between species resulting in different disposition and PD characteristics; in such instances a comprehensive PK/PD modeling proves extremely useful in projecting effects and side effects in humans as a method of inter-species extrapolation.

Another reason for using population PK approach is to obviate the inherent weaknesses built into pure PK studies. These PK studies are designed such as to limit the variation between and among individuals through strict acceptance criteria of healthy subjects; historically, kineticists have tried to define PK parameters within a very narrow range as a measure of goodness of a study; the reality is that when a drug is used by consumers, it will be subject to all types of variations possible, including many that could not have been envisioned in the development phases. Therefore, focusing on a single PK parameter (such as half-life or clearance) on determining appropriated dose is not always the best choice.

In contrast to traditional PK evaluation, the population PK approach encompasses the following information:

- The collection of relevant PK information in patients who are representative of the target population to be treated with the drug.
- The identification and measurement of variability during drug development and evaluation.
- The explanation of variability by identifying factors of demographic, pathophysiologic, environmental, or concomitant drug-related origin that may influence the PK behavior of a drug.
- The quantitative estimation of the magnitude of the unexplained variability in the patient population.

Population PK modeling is also used to estimate population parameters of a "response surface model" in Phase I and late Phase IIb clinical drug development. It is this stage when the developer decides on what is the most appropriate application of the drug. Since the samples at this stage can be extensive, there is little need to resort to complex data analysis systems and the use of the two-stage method often suffices. The more complex nonlinear, mixed-effects model can be used when it is desired to pool the data from several studies. Population PK modeling is also used in early Phase IIa and Phase III drug development to gain information on drug safety (vis-à-vis efficacy) and to gather additional information on drug PK in special populations, such as the elderly or pediatric populations. This approach is also useful in postmarketing surveillance (Phase IV) studies. The studies performed during Phase III and Phase IV clinical drug development allow use of a full-population sampling PK study-design (few blood samples drawn from several subjects at various time points). This sampling design can provide important information during new drug evaluation, regulatory decision-making, and drug labeling.

There are two common methods for obtaining estimates of the fixed effects (the mean) and the variability: the two-stage approach and the nonlinear, mixed-effects modeling approach. The two-stage approach involves multiple measurements on each subject. The nonlinear, mixed-effects model can be used in situations where extensive measurements cannot or will not be made on all or any of the subjects.

When properly applied, population PK studies in patients combined with suitable mathematical and statistical analyses (e.g., using nonlinear, mixed-effects modeling) are valid, and on some occasions, are preferred alternatives to extensive studies; this applies to both PK and PD studies. When large data are not available, the traditional two-stage approach is not applicable because estimates of individual parameters cannot be obtained; in such instances, nonlinear, mixed-effects modeling proves valuable. This modeling approach is developed from the realization that where PK and PD are to be investigated in patients, practical considerations should prevail in determining how much data can be gathered. This would inevitably have to be less stringent because the patients subjected to dose-determining studies are less likely to be tolerant than the healthy subjects. Other approaches such as the naive averaged-data modeling can also be used which are also simpler systems; in this model, the study of a population sample rather than that of an individual is used as a unit of analysis for the estimation of the distribution of parameters and their relationships with covariates within the population. This approach uses individual PK data of an observational (experimental) type, which can be sporadic, unbalanced, or fragmentary, compared to extensive uniform data in well-designed PK studies. Analysis according to the nonlinear, mixed-effects model in such situations provides estimates of population characteristics that define the population distribution of the PK (or PD) parameter. In the mixed-effects modeling context, the collection of population

characteristics is composed of population mean values (derived from fixed-effects parameters) and their variability within the population (generally the variance and covariance values derived from random-effects parameters). A nonlinear, mixed-effects modeling approach to the population analysis of PK data, therefore, consists of estimating directly the parameters of the population from the full set of individual concentration values. The individuality of each subject is maintained and accounted for, even when data are not extensive.

18.7.6 Pharmacokinetics in Disease States

18.7.6.1 Hepatic Impairment

The time course of drug concentration (and hence, in most instances, its PD) is determined by the mechanisms that remove the drug from the body, primarily the liver and the kidney. The liver clears drugs through a variety of oxidative and conjugative metabolic pathways or through biliary excretion of unchanged drugs or metabolites. Alterations of these excretory and metabolic activities by hepatic impairment can lead to drug accumulation or, less often, failure to form an active metabolite. Kidneys remove drugs by filtration and secretion.

It is a well-established fact that hepatic disease alters the absorption and disposition of drugs (as described in PK studies) as well as their efficacy and safety (as described in PD studies). The type of studies that have established this fact include common hepatic diseases, such as alcoholic liver disease and chronic infections with hepatitis viruses B and C, and less common diseases, such as acute hepatitis D or E, primary biliary cirrhosis, primary sclerosing cholangitis, and α-antitrypsin deficiency. Liver disease can also alter kidney function, which can lead to accumulation of a drug and its metabolites, even when the liver is not primarily responsible for elimination. Liver disease may also alter PD effects (such as the increased encephalopathy with certain drugs in patients with hepatic failure). The specific effect of any disease on hepatic function is often poorly described and highly variable, particularly with regard to effects on the PK and PD of a drug.

Hepatic function is assessed using markers such as the endogenous substances (e.g., bilirubin and albumin) affected by the liver; or by using functional measures such as prothrombin time; or by using the ability of the liver to eliminate marker substrates, such as antipyrine,[33] indocyanine green,[33] monoethylglycine-xylidide,[34] or galactose.[35] Clinical variables have also been studied — including ascites or encephalopathy, nutritional status, peripheral edema, and histologic evidence of fibrosis, and combinations of variables, such as the Child-Pugh classification for alcoholic cirrhosis and portal hypertension,[36,37] the Mayo risk scores for primary biliary cirrhosis and primary sclerosing cholangitis,[38] and the Maddrey-Carithers discriminant function for acute alcoholic hepatitis.[39,40] Despite extensive efforts, no single measure or group of measures is adequate to describe the impact of renal impairment universally; in almost all instances, the sponsor will need to develop a detailed set of parameters to adequately describe the PK and PD in hepatic impairment. FDA and other regulatory agencies recommend PK studies in patients with impaired hepatic function if hepatic metabolism or excretion accounts for a substantial portion (greater than 20% of the absorbed drug) of the elimination of a parent drug or active metabolite. Even if this 20% rule of thumb does not apply, the FDA requires liver impairment studies for drugs (or their metabolites) with narrow therapeutic or toxic index. This is necessary to take into account population variation in hepatic function. It is noteworthy that unless the sponsor shows that the drug is not extensively metabolized, it is assumed to be removed primarily from the liver and thus requiring liver impairment studies.

For some drugs, hepatic functional impairment is not likely to alter PK sufficiently to require dosage adjustment. In such cases, a study to confirm the prediction is generally not important. Drug properties that support this conclusion include drugs excreted entirely via renal routes, with no involvement of the liver, drugs metabolized in the liver to a small extent (less than 20%), and drugs in which the therapeutic range is wide, so that modest impairment of hepatic clearance will not lead to drug toxicity directly or by increasing its interaction with other drugs. This also applies to gaseous or volatile drugs that are primarily eliminated via the lungs. For drugs intended for single-dose administration, hepatic-impairment studies are not generally useful, unless clinical concerns suggest otherwise.

Population PK screening in Phase II and Phase III is useful in assessing the impact of altered hepatic function (as a covariate) in PKs, if those patients are not excluded from Phase II and III trials, and if there is sufficient PK information collected about the patients to characterize them reasonably well. If a population PK approach is used, patients in Phase II and III studies are assessed for encephalopathy, ascites, serum bilirubin, serum albumin, and prothrombin time (which are components of the Child-Pugh score) or a similar group of measures of hepatic function. The population PK study, then, would include the following features:

- Preplanned analysis of the effect of hepatic impairment
- Appropriate evaluation of the severity of liver disease
- A sufficient number of patients and a sufficient representation of the entire range of hepatic functions to allow the study to detect PK differences large enough to warrant dosage adjustment
- Measurement of the unbound concentrations of the drug, when appropriate
- Measurement of the parent drug and the active metabolites

These features are important if the sponsor intends to use the results to support a conclusion that no dosage adjustment is required for patients with impaired hepatic function. Pharmacodynamic assessments may be useful in studies designed to assess the effect of altered liver function, especially if concentration–response data are not available or if there is a concern that an altered hepatic function could alter the PD response.

Plasma-concentration data (and urine-concentration data, if collected) are analyzed to estimate the measures or parameters — such as AUC, peak concentration (C_{max}), apparent clearance (CL/F), renal and nonrenal clearance, apparent volume of distribution (Vdz or $Vdss$), and terminal half-life ($t_{1/2}$) — that describe the PKs of the drug and its active metabolites. Where relevant, measures or parameters can be expressed in terms of unbound concentrations, such as the apparent clearance relative to the unbound drug concentration (CL_u/F = dose/AUC_u, where subscript u indicates unbound drug). Noncompartmental or compartmental modeling approaches to parameter estimates can be used.

Relationships between hepatic functional abnormalities (e.g., hepatic blood flow, serum albumin concentration, or prothrombin time), overall impairment scores (such as the Child-Pugh), and selected PK parameters (such as total body clearance, oral clearance, apparent volume of distribution, unbound clearance, or dose-normalized area under the unbound, concentration–time curve) are sought using linear and nonlinear models. A regression approach for continuous variables describing hepatic impairment and PK parameters is appropriate with the understanding that some correlations will rely on categorical variables (such as the Child-Pugh). Typically, modeling results include parameter estimates of the chosen model and measures of their precision (standard errors or confidence intervals). Prediction-error estimates are also desirable to assess the appropriateness of the model.

18.7.6.2 Renal Impairment

Similar to the recommendations made above in adjusting the dosing in hepatic function modification, a PK study in patients with impaired renal function is recommended if the drug is significantly removed from the body through kidneys, and therefore the PK parameters are likely to be different in patients with impaired renal function. Generally, drugs given in a single dose, drugs with high therapeutic indices and eliminated via hepatic clearance, or gaseous or volatile drugs eliminated through the lungs may not require studies in renal impairment. This rule-of-thumb changes, however, where renal impairment affects hepatic metabolism or where the impact of dialysis on the PK of a drug should be considered. Note that this study of changes in disposition function in renal impairment affects both the active drug and its metabolites, and it is particularly important where a narrow therapeutic or toxic index is observed. In some instances, a study in patients with renal impairment is required even though the drug is not significantly removed by the kidneys but where compromised renal function affects plasma protein binding of drugs. In such instances, a drug highly cleared by liver (but dependent on protein binding) may show a significant difference in its disposition characteristics in renal impairment. Patients on dialysis often require larger doses; dialysis often significantly affects the PK requiring dose adjustment. As a result for patients with end-stage renal disease requiring dialysis, PK should be studied under both dialysis and nondialysis. If the drug and metabolites have a large unbound volume of distribution (V), only a fraction of the amount in the body will be removed by dialysis. Measurements, such as plasma creatinine or creatinine clearance, have been used successfully to adjust dosing regimens for drugs eliminated primarily by the kidneys.

In some cases, renal impairment studies are conducted to establish the thesis that a dosage adjustment is not necessary; these studies certainly make an impressive presentation to regulatory authorities and more than that allow the sponsor to make certain additional claims on the label. Note that renal impairment is a major pathologic change that induces many physiologic changes and unrelated effects are often observed unexpectedly, such as altered absorption and distribution of drugs in renal impairment, even when the renal route does not represent a significant elimination mechanism. The kidneys also participate in the metabolism of drugs, and produce hormones that can indirectly affect pathophysiology of disease and drug disposition. For example, in renal impairment, secretion of erythropoietin is reduced leading to anemia; were a drug extensively bound to components of red blood cells, its disposition is likely to be altered in renal impairment.

18.7.6.3 Pharmacokinetic Studies in Special Populations

FDA places special importance on PK studies in pediatric patients because of the essential differences in physiological characteristics between adults and children; variance can sometimes be substantial. In pediatric populations, growth, and developmental changes affect the rates of adsorption, distribution, metabolism, and excretion (ADME) and lead to changes in PK measures and parameters. To achieve AUC and C_{max} values in children similar to values associated with effectiveness and safety in adults, it may be important to evaluate the PK of a drug over the entire pediatric age range in which the drug will be used. Where growth and development are rapid, adjustment in dose within a single patient over time may be important to maintain stable systemic exposure.

The PK in the elderly is another subject of great interest to regulatory authorities. Although in most instances, unlike pediatric PKs, some linear extrapolations can be drawn on the basis of such parameters as lean body mass and age; these should, nevertheless, be studies and validated.

18.7.7 Computational Support

During the past few years, computational efficiency has advanced, and many software programs have become available — often on free domains — that are useful in the simulation of PK models. The following "Software for Use in Pharmacokinetic/Pharmacodynamic Modeling" sidebar lists some of these programs:

GENERAL-PURPOSE HIGH-LEVEL SCIENTIFIC COMPUTING SOFTWARE

Berkeley Madonna, University of California at Berkeley, http://www.berkeleymadonna.com
MATLAB-Simulink, The MathWorks, Inc., http://www.mathworks.com/
MLAB, Civilized Software, Inc., http://www.civilized.com
GNU Octave, University of Wisconsin, http://www.octave.org/

BIOMATHEMATICAL MODELING SOFTWARE

ADAPT II, Biomedical Simulations Resource, USC http://bmsr.usc.edu/
ModelMaker, ModelKinetix, http://www.modelkinetix.com/
NONMEM, University of California at San Francisco and Globomax Service Group, http://www.globomaxservice.com/
Stella, High Performance Systems, Inc., http://www.hps-inc.com/
WinNonlin, Pharsight Corp., http://www.pharsight.com
SAAM II, SAAM Institute, Inc., http://www.saam.com

TOXICOKINETIC SOFTWARE

ACSL Toxicology Toolkit, AEgis Technologies Group, Inc., http://www.aegistg.com/

PHYSIOLOGICALLY BASED CUSTOM-DESIGNED SOFTWARE

GastroPlus, Simulations Plus, Inc., http://www.simulations-plus.com
Pathway Prism, Physiome Sciences, Inc., http://www.physiome.com
Physiolab, Entelos, Inc., http://www.entelos.com/

One should definitely examine them to appreciate how easy it has become to create and analyze PK and PD models. If you are interested in further information, a literature search on of electronic databases for material on PK/PD modeling should use both the spellings "modeling" and "modelling" because they are used interchangeably.

18.8 Concentration–Response Relationships

The basic model for ascertaining dose–effect relationships is derived from the sigmoidal effect model, which correlates maximal response (E_{max}), placebo (E_0), and the dose producing 50% of effect (ED_{50}):

$$\text{Effect} = E_0 + \frac{E_{max} \times \text{dose}}{ED_{50} + \text{dose}} \tag{18.16}$$

The effect is related to PK characteristics through the clearance model wherein the ED_{50} is correlated to the plasma concentration, CL, and bioavailability; note that EC_{50} is the one-half of the concentration found at E_{max}:

$$ED_{50} = \frac{\text{CL} \times EC_{50}}{\text{Bioavailability}} \tag{18.17}$$

The significance of the term EC_{50}, a mathematical parameter only, lies in its independence of any PK or dosage-form factors and depends only the type of drug used; note that ED_{50} on the other hand is highly dependent on the delivery system. During the development

phase, a drug company may study several dosage forms which require development of ED_{50} values for each; on the other hand, EC_{50} is a more fundamental parameter related to potency and toxicity of drug that can be readily obtained from single-dose studies wherein the complete plasma level profile provides essential information; this is in sharp contrast to dose-range studies where in at least three doses are administered.

The usual definitions of maximal effect (E_{max}) and potency (EC_{50} or IC_{50}) require another look at this stage before proceeding to developing mathematical relationship between drug concentration, dose, and body functions using a PK/PD approach.[41]

Potency comprises both action and inhibition of action and is predicted by the Hill model; though a 50% level is chosen, it is an arbitrary percentage and other values such as 60 or 40% action can also be calculated and used. Potency is not a relevant factor unless it is so low that the dose requirement is very high (to a level where nonlinear binding with albumin can be observed, resulting in nonlinear kinetics) or where the serious side effects are dose-dependent and make an effective dose unacceptably toxic. The potency, EC_{50}, is expressed in a mechanistic equilibrium model where the action is direct:

$$EC_{50} = \frac{K_d}{(1+\tau)} \tag{18.18}$$

where K_d is drug affinity and τ a transducer constant described as

$$\tau = R_{total}/KE \tag{18.19}$$

Here R_{total} is the size of receptor pool and KE is the dissociation constant for the agonist–receptor complex formed; it is a measure of drug-binding properties and physiologic-system response. Note that low KE means a highly efficient stimulus–response relationship (which will yield a high value of τ) and thus the term EC_{50} or potency turns out to be dependent on the size of receptor pool, affinity, and the efficiency of binding. How EC_{50} compared to K_d determines whether a drug is efficacy- or affinity-driven agonist. When EC_{50} is much lower than K_d, the drug is efficacy-driven, since in this case $\tau > 1$; an opposite case applies where τ is close to 1 when EC_{50} is lower than K_d (though not too low). Efficacy-driven drugs are little affected when the density of receptor is altered, meaning that there is a change in potency but not in the maximal effect. The affinity-driven drugs, on the other hand, are affected by changes in receptor density that results in changes in maximal response. The latter category of drugs have complicated PK/PD relationship and is often difficult to model since highly efficacious agonists are likely to produce maximal response regardless of tissue or species studied, and low-efficacy drugs of this type may interact with only specific tissues and are more likely to show interspecies differences, something that is of great significance in scaling of pharmacologic and clinical effect from one species to another. A lot of work is done on developing good animal species models in the early drug development; an early appreciation whether a drug is likely to have larger interspecies differences is crucial in reducing the cost of this stage of drug development.

When studying the maximal efficacy (E_{max}), one should not confuse it with potency or intrinsic effects; two drugs with entirely different potency can elicit the same maximal response, except that the dose will be different. This parameter is often difficult or impossible to study experimentally as increasing the dose invariably introduces safety issues. Also, the maximal effect is what is observed and not necessarily what an organ is capable of reaching. The maximal effect (E_{max}) of system

$$E_{max,\,drug} = \frac{E_{max,\,system}\,\tau}{\tau+1} \tag{18.20}$$

is of great importance in PK/PD modeling; Toutain[41] provides an excellent review of these equations (described below). The classical approach in dose selection is dose ranging, which is based on a parallel dose–response design, where subject are randomly assigned to dose levels and the response analyzed using statistical model where the observed response (Y_{ij}) is in ith subject and jth dose, θ_{ij} is the mean response from jth dose, and ε_{ij} is the error in the observed vs. the expected response from the jth dose:

$$Y_{ij} = \theta_j + \varepsilon_{ij} \tag{18.21}$$

In the crossover design, there is a drawback because it is unable to provide information about the shape of the individual dose–response curve. In a crossover design, the "effective dose" is determined by statistical analysis (by testing the null hypothesis) and, therefore, is highly dependent on the power of the study. This results in crossover studies in which sample sizes are small and the errors are large, and which end up with recommendations for higher doses — something the drug companies actually prefer because it is always easy to reduce dosage later without additional, expensive trials. A combination of these two approaches, i.e., dose-range and crossover provides greater robustness to the data obtained.

Besides the continuous effect models, drugs are often fitted to all-or-none models where the responses such as disappearance of arrhythmia in response to a drug dose cannot be graded; the same results hold when we classify the response ending in cure or not, or presence or absence of a given side effect or effects. In such instances, EC_{50} is the median concentration for which half of the subject population is above the threshold and the slope of the curve becomes the variance of the threshold in that population.

Graphic representation of cumulative frequency distribution of selected effects as a function of concentration is also prepared where ratio of median effective concentration is considered for drug selectivity using different endpoints.

18.9 Pharmacodynamic Models

The PD models fall under two categories: graded or quantal of fixed-effect model. Graded refers to a continuous response at different concentrations, whereas the quantal model would evaluate discrete response such as dead or alive, desired or undesired and are almost invariably clinical end points.

18.9.1 The Hill Model

In graded response one can resort to using surrogates, and the classic Hill model (or sigmoid model as described in Equation 18.16 above) is used to correlate observed effect ($E(t)$) with the concentration modified by an exponent that is called Hill coefficient; in classic sigmoidal model, the Hill coefficient (h) would be equal to 1:

$$E(t) = E_0 + \frac{E_{\max} \times C^h(t)}{EC_{50}^h + \left[C^k(t)\right]} \tag{18.22}$$

The shape of curve is modulated by varying the Hill coefficient to fit the observed behavior. Sensitivity of a drug is how concentration translates into effect as described by the shape coefficient or the Hill coefficient; at a value of 1, it is a classical parabola. E_0 is

generally the baseline or placebo effect; in the event when the response involves an inhibitory effect (IC_{50}), the response is subtracted from observed response and thus the above equation reads as follows:

$$E(t) = E_0 - \frac{E_{\max} \times C^h(t)}{IC_{50}^h + C^h(t)} \tag{18.23}$$

where there is full inhibition possible, the term E_0 is replaced by $E_{\max}$, reducing the above equation to a fractional equation:

$$E(t) = E_{\max}\left[1 - \frac{c^H(t)}{IC_{50}^h + C^h(t)}\right] \tag{18.24}$$

where the term in parenthesis is the fraction of maximum effect remaining at any given time and ranges from 1 to 0; the fraction of reached is given by the term subtracted from 1. This is therefore labeled as fractional Hill model.

The models presented above in Equations 18.23 and 18.24 have three types of parameters: the independent parameter of concentration, the dependent parameter of effect (E), and those parameters that are obtained from the observed relationship between E and C: E_0, $E_{\max}$, EC_{50} or IC_{50} or h.[42,43] The variation of these parameters is pivotal in establishing the relationship between PD and PK properties. When evaluating this relationship of parameters, it is important to differentiate drug action and drug response. The action pertains to mechanism of action and the response or effect, the outcome. For example, the action of aspirin is to block cylcooxygenase, and the effect is reduction in atheromas. Where the intent is to determine a dosing regimen, the response or effect is more important; however, there instances where it is more relevant to study drug levels at the site of action because a direct relationship is not always established between these levels and the response. As an example, the levels of NSAIDs at the site of action can be a useful parameter since the response is highly variable. The term $C(t)$ in Equations 18.22 to 18.24 is an independent parameter that represents blood or plasma levels; however, this can well be urinary concentration or concentration in other biological tissue samples.[44,45] Instances where an instantaneous equilibration between plasma concentration and the biophase is possible such as in the case of angiotensin-converting enzyme (ACE), where equilibration is quickly achieved between plasma and surface of blood vessels, Equation 18.18 in monocompartment situation would be as follows (instantaneous equilibration model):

$$E(t) = E_0 + \frac{E_{\max}\left[\frac{\text{Dose}}{Vc}\exp\left(-\frac{\text{CL}}{Vc} \times \text{time}\right)\right]^h}{EC_{50}^h + \left[\frac{\text{Dose}}{Vc}\exp\left(-\frac{\text{CL}}{Vc} \times \text{time}\right)\right]^h} \tag{18.25}$$

Notice how Equation 18.25 correlates Equation 18.16, which represents PK/PD parameters while Equation 18.25 is a PK parameter-dependent equation. Equation 18.25 introduces another independent parameter, time, allowing determination of both the optimal dose as well as optimal dosing interval for drugs intended for multiple dosing such as antibiotics, antihypertensive drugs, and antiepileptic drugs.

Equations 18.22 and 18.25 are intended for drugs where a relationship between drug concentration and effect exists and for these drugs the profiles are meaningful; however, when concentration is of lesser importance such as where a drug induces a long-term toxicity, a better approach is to use AUC or the time when a threshold concentration is achieved is more important.[46] This can be studied by integrating Equations 18.22 and 18.25 to give a relationship between E and AUC (the independent variable, determined by $F \times \text{Dose} \times \text{CL}$; AUC_{50} is the exposure that yields $E_{max}/2$) in a Hill model:

$$E = E_0 + \frac{E_{max} \times (\text{AUC})^h}{\text{AUC}_{50}^h + (\text{AUC})^h} \tag{18.26}$$

In those instances of time-dependent drugs where concentration is related to a direct but unobserved action, A_{direct}, the observed response is related to the cumulative action of drug through this direct (and primary) action:

$$E_{observed} = \frac{E_{observed,\,max} \times \text{AUC} - A_{direct}}{\text{AUC} - A_{direct,\,50} + \text{AUC} - A_{direct}} \tag{18.27}$$

The term $\text{AUC} - A_{direct,50}$ is the duration of maximum direct action that produces half the maximal observed response. Where the concentration is much smaller than EC_{50}, dividing the above equation by EC_{50} gives

$$E_{observed} = \frac{E_{observed,\,max} \times (\text{AUC}/EC_{50})}{(\text{AUC}/EC_{50})_{50} + (\text{AUC}/EC_{50})} \tag{18.28}$$

Now the term AUC/EC_{50} becomes the independent potency parameter that will predict efficacy of concentration-dependent drugs such as antibiotics (ala MIC_{50} or MIC_{90} that predict ordinal effects), and the term $(\text{AUC}/EC_{50})_{50}$ has dimensions of time.

18.9.2 The Receptor Theory Model

The Hill coefficient measures system cooperativity in receptor theory models where $h>1$ means that agonist molecules facilitate binding of subsequent model arriving at receptor. For drugs with low h value, the drug effects are only moderate over a wide range of dosing meaning that the curve is shallow giving measurable responses at low plasma levels; in such cases the terminal half-life proves very useful in predicting the duration of effect.[47] On the other hand, where the slope is steep, variations in concentration become more important and around EC_{50} these variations can produce a wide range of effects, from no effect to maximal response. This consideration is more important where the drug has low therapeutic index. Whereas antibiotics can be classified into high and low h using the Hill model, this does not necessarily mean that one class is time-dependent and the other concentration-dependent. More appropriately, PK properties and the concentration–effect profile are integrated where the effect becomes proportionate to time above a critical concentration or the AUC; at a high h value of above 5, the concentration range ends up in a threshold and graded PD model turns into a quantal model.[48]

The dose–response relationship in a quantal model can be analyzed with the help of a logistic model where we calculate probability of an event at a given concentration or AUC or dose:

$$\pi_{(\text{outcome})} = \frac{e^{\alpha+\beta x}}{1+e^{\alpha+\beta x}} \tag{18.29}$$

The exponent α fixes location while β is the slope; the probability of an event not happening is simply

$$1-\pi_{(\text{outcome})}=\frac{1}{1+e^{\alpha+\beta x}} \tag{18.30}$$

A ratio of odds whether an event is likely or not is simply the ratio of odds of happening and odds of not happening:

$$\frac{\pi_{(\text{outcome})}}{1-\pi_{(\text{outcome})}}=e^{\alpha+\beta x} \tag{18.31}$$

The logit (L) is the natural logarithm given by

$$L=\alpha+\beta x \tag{18.32}$$

The logistic model differs from Hill model because in the logistic model the dependent variable is categorical, though both models use the same parameters and the same underlying model. In Hill model, the probability of outcome can be calculated from the known concentration C and the unknown median concentration EC_{50}; when EC_{50} (the concentration at which the probability of effect is 50%) is very low, the probability of action is always certain, almost equal to 1 but when EC_{50} is very high compared to C then it is simply the ratio of C and EC_{50}:

$$\pi_{(\text{outcome})}=\frac{C^h}{EC_{50}^h+C^h} \tag{18.33}$$

Comparison of Equation 18.33 with Equation 18.29 shows that α is $\ln(EC_{50})^h$, x is $\ln C$, and β corresponds to the Hill coefficient, h, which now represents interpatient variability.

18.9.3 The Hysteresis Loop Phenomenon

If the effect is proportional to concentration at all ranges and times, then we will have a simple relationship with effect rising with concentration and the reaching a plateau, theoretically. The same type of relationship is always observed if we were to sample the drug at the site of action. In real situations, this seldom happens for two reasons. First, the PK profile need not be similar to PD model due to a lag in drug distribution, for example; and second, the drug response on the rising curve of concentration may not be the same as in the falling range. These two factors lead to a plot if effect is plotted against concentration that appears like a loop, a hystersis loop; if an inverse relationship exists in the suppression of an effect, the loop will go clockwise, and in rare administration of drug via arterial route, the site of action equilibrates faster than the plasma concentration or where there is an accumulation of antagonist, the loop takes the form of proteresis.[47] Several interesting correlations are drawn from appearance of hystersis loop phenomenon. If this phenomenon is observed, then an attempt should be made to take into account the delay in action through mathematical intervention such as volume of distribution estimation since biophase is poorly perfused or protected by a barrier[48] and developing understanding of the underlying mechanism such as induction of an enzyme, conversion of drug to an active form, etc. Hystersis is also observed when over time the tissue becomes sensitized thereby giving higher response. When there are changes in the neurotransmitters (increased or

decreased) or other such actions where the drug action is indirect, it is better to employ indirect response models.

Therefore, we can conclude that hystersis can be of PK origin or of PD origin; in the case of latter origin it is desired to sample the site of action if possible, otherwise resort to hypothetical effect-compartment modeling (see below). Running numerous steady-state experiments does not prove very useful or practical in such situations.

18.9.4 The Effect-Compartment Model

The milieu of effect can be represented by a hypothetical compartment, as if this were a continuous space or a composite of similar spaces; the concentration within the effect compartment can then be related to the observed effect:

$$E(t) = \frac{E_{\max} \times Ce(t)}{EC_{50} + [Ce(t)]} \tag{18.34}$$

It is important to understand that $Ce(t)$ is not a measurable concentration but it is dependent on the distribution of drug from the central compartment (where the drug is administered) and any delays in reaching a certain level of $Ce(t)$ are attributed to this distribution delay; similarly the terms $E_{\max}$ and EC_{50} can only be obtained indirectly. Since the rise in $Ce(t)$ is dependent on the amount of drug entering the effect compartment, the equilibration constant between the effect compartment and the central compartment controls this function:

$$\mathrm{d}Ae/\mathrm{d}t = (K_{1e} \times A_1) - (Ke_0 \times Ae) \tag{18.35}$$

The input function is the product of amount in the central compartment A_1 and the entry rate constant K_{1e}; the output function is given by the amount in the effect compartment Ae and the outward rate constant Ke_0. Both rate constants are of first order. In most instances $Ce(t)$ or the amount in the effect compartment is much smaller compared to the amount in the central compartment and thus a small error assumption will assume that the drug is eliminated directly from the effect compartment rather than first returning to the central compartment, which, in reality is the case. This assumption allows simplification of the above equation. Generally, the concentration in the effect compartment is easily described in the traditional PK modeling situations. For example, in a classic single-compartment model

$$Ce(t) = \frac{K_{1e} \times \mathrm{Dose}}{V_{\mathrm{e}}(Ke_0 - K_{10})}\left(\mathrm{e}^{-K_{1e}t} - \mathrm{e}^{Ke_0 t}\right) \tag{18.36}$$

The term V_{e} refers to the distribution volume of the effect compartment and thus the effect compartment becomes synonymous with the central compartment. The above equation contains three unknown parameters making it impossible to predict the concentration or the effect ($E(t)$) as a function of time; if we define the ratio of concentration in the effect compartment and the central compartment, or the partition coefficient, at equilibrium,

$$K_{\mathrm{p}} = K_{1e}V_1/Ke_0V_{\mathrm{e}} \tag{18.37}$$

then this will allow rewriting the above equation as

$$Ce(t) = \frac{Ke_0 \times \mathrm{Dose} \times K_{\mathrm{p}}}{V_1(Ke_0 - K_{10})}\left(\mathrm{e}^{-K_{10}t} - \mathrm{e}^{-Ke_0 t}\right) \tag{18.38}$$

In this equation, we now have the term for the volume of central compartment, V_1, instead of the volume term in the effect compartment. If we use the term $Ce(t)/K_p$ rather than $Ce(t)$ in the effect model, we can rewrite the effect as a function of time:

$$E(t) = \frac{E_{max} \times [Ce(t)/K_p]}{(EC_{50}/K_p) + [Ce(t)/K_p]} \tag{18.39}$$

By substituting the value of $Ce(t)$, we get

$$E(t) = \frac{E_{max} \cdot \dfrac{Ke_0 \times \text{Dose}}{V_1(Ke_0 - K_{10})}\left(e^{-K_{10}t} - e^{-Ke_0 t}\right)}{(EC_{50}/K_p) + \left[\dfrac{Ke_0 \times \text{Dose}}{V_1(Ke_0 - K_{10})}\left(e^{-K_{10}t} - e^{-Ke_0 t}\right)\right]} \tag{18.40}$$

Now we have an equation that can be resolved from PD data over time and having on hand the basic PK parameters. It should be understood that the term EC_{50}/K_p is the EC_{50} that produces $E_{max}/2$ under equilibrium condition; this parameter is of practical use in adjusting dosing regimen. The K_p factor can only be estimated by sampling the biophase, which is generally not possible; in such instance, the plasma concentration serves a clinical end point.

The equilibration of drug between plasma and biophase is determined by the parameter Ke_0 and the equilibration half-time $\ln 2/Ke_0$, and $T_1/2Ke_0$ is the length of time it takes for plasma concentration and effective concentration to reach equilibrium; this time can range from a few minutes to several hours. For drugs with short half-time, the value of Ke_0 is large and an equilibration is reached fast, and thus the plasma concentration is a good indicator of biophase levels; for drugs with low Ke_0 (lower than the terminal elimination-rate constant), an equilibrium will never be reached. The time at which peak concentration is achieved (Te_{max}) is expressed as

$$Te_{max} = \frac{\ln\left(\dfrac{Ke_0}{K_{10}}\right)}{Ke_0 - K_{10}} \tag{18.41}$$

The plasma concentration at Te_{max} when divided by the dose gives an estimate of distribution volume that can be used to calculate dosing for fast equilibrating drug, in particular, such as anesthetics. This distribution volume term often proves more useful than the traditional central volume of distribution or the steady-state distribution volume.

The term Ke_0 determines how a hysteresis loop closes and thus represents a steady-state-effect relationship.[49,50] Using this parameter, there is no need to predefine a PD model as the nature of the model is decided by the hysteresis loop encountered. One of the most common mistakes encountered in PD modeling is to set E_{max} to 100%. At those instances where the concentration profiles are not well-defined or cannot be defined by traditional PKs such as when modeling an endogenous compound, e.g., insulin, erythropoietin, etc., administered in specialized dosage forms, one can model the observed concentration into PD models by smoothing the function using traditional linear or cubic-spline functions.[51]

18.9.5 The Indirect Action Models

When the concentration of a drug cannot be directly related to effect, indirect modeling is suggested. This will most aptly apply to situations where, for example, the response is not directly a result of action (e.g., binding a receptor site) as there may be several steps involved in between the action and response with their own specific mathematical relationships. Unlike the direct action models where any delay in the response is likely a result of PK phenomenon, the delay where indirect models are used is a result of intrinsic nature of drug action and response relationship. Several models can be used in such instances:

$$\frac{dR}{dt} = K_{in} - (K_{out}R) \tag{18.42}$$

The rate equation describes the variation in response variable R (with initial value of R_0); the measure response appears at a constant rate (zero order) of K_{in} and is eliminated by K_{out}, the first-order constant. The indirect response models will generally fall into two categories: inhibition or stimulation function. The inhibition response is classically described in terms of IC_{50}, the drug concentration that produce 50% of maximal inhibition (e.g., action of synthetic glucocorticoid on adrenal glands or effect of furosemide on sodium absorption in the loop of Henley) and I_{max} is a number from 0 to 1 where 1 represents total inhibition:

$$I(t) = -\frac{I_{max} \times C(t)}{IC_{50} + [C(t)]} \tag{18.43}$$

Substituting this equation in the above model gives

$$\frac{dR}{dt} = K_{in}\left\{1 - \frac{I_{max} \times C(t)}{IC_{50} + [C(t)]}\right\} - (K_{out} \times R) \tag{18.44}$$

Similarly, where a stimulation action (e.g., antipyretic effect of NSAIDS through thermolysis or production of cAMP by bronchodilator beta-2 agonist[52]) is involved, the stimulation response

$$S(t) = \frac{S_{max} \times C(t)}{SC_{50} + [C(t)]} \tag{18.45}$$

when substituted into PD model gives

$$\frac{dR}{dt} = K_{in}\left\{1 + \frac{S_{max} \times C(t)}{SC_{50} + [C(t)]}\right\} - (K_{out} \times R) \tag{18.46}$$

The maximal effect in the models given above occurs later than the time when C_{max} is reached[53] since the drug produces an incremental effect, either inhibition of stimulation, provided the concentration remains above IC_{50} or SC_{50}. How the response returns to its baseline after concentration decay is governed by K_{in} the zero-order input rate and the rate

TABLE 18.3

Examples in Literature Where PK/PD Studies Have Been Reported

Drug Reference	Findings
Apomorphine:[56] stimulated growth hormone release via dopamine D2 receptors.	$C(t)$ modeled according to two-compartment model with zero-order and first-order absorption Pharmacokinetic/pharmacodynamic relationship modeled using E_{max} Hill model with first-order absorption. Modeled parameters matched experimental parameters when bicompartmental model with zero-order input was used. Linear PKs, anticlockwise hysteresis loop established for all doses studied. Apomorphine and growth hormone concentration predicted with good accuracy
Erythropoietin recombinant[57]	Pharmacokinetic model with negative feedback loop for endogenous erythropoietin; PD analysis using changes in reticulocyte count, serum level of soluble transferring receptors, soluble transferring receptors-serum protein ratio as index of therapeutic efficacy. Flip-flop PK model and PD relationship based on sigmoid E_{max} model related markers to erythropoietin dosing; PK/PD modeling predicted delay in response using half-life
Digoxin[58]	Pharmacokinetic/pharmacodynamic model using nonlinear, mixed-effects model in two compartment, best described time course of concentration; strong correlation with creatinine clearance; predicted concentration at the effect site and in reduction of heart rate during atrial fibrillation using population kinetic approach
TF-505:[59] nonsteroidal and noncompetitive inhibitor of type I and II 5α-reductases.	Pharmacokinetic/pharmacodynamic evaluated using indirect response modeling with inhibition of input
Mefloquine[60]	The *Plasmodium falciparum* malaria PD model successfully described the antimalarial effect of artemisinin, mefloquine, and a combination of the two drugs
Telithromycin[61]	Pharmacokinetic/pharmacodynamic modeling indicates that the AUC/MIC ratio is predictive of outcome for this antibacterial suggesting that telithromycin is suitable for short-duration therapy and a once-daily dosing regimen even against β-lactam and macrolide-resistant pathogens
Insulin[62]	Two-compartment PK model; the indirect PD response model is a more appropriate approach for modeling the PK/PD of insulin than the effect-compartment link model
Angiotensin II receptor antagonist class[6]	A common population PK model (two compartment, mixed zero- and first-order absorption, two metabolite compartments) was applied. A standard, nonlinear, mixed-effect modeling approach was used to estimate the drug-specific parameters and their variability. This model can be used to further refine the use of this class of drugs
Insulin[64]	Bioavailability of human insulin assessed from pharmacological data. After extravascular administration only the time course governs the observed pharmacological effects. The pharmacological data was translated into theoretical plasma-concentration data using the PK/PD model. The results of the PK/PD analysis indicate that the doses administered can be accurately predicted from pharmacological data
Octreotide acetate[65]	Population PK analysis following a single dose of 30 mg microencapsulated octreotide acetate intramuscularly; the PK profile of microencapsulated octreotide acetate was effectively described by the derived population model. The relationship between IGF-1 and drug concentration could be used to guide optimization of therapeutic octreotide dosage regimens
Thalidomide[66]	A four-compartment model consisting of a two-compartment model for each enantiomer, with elimination from both compartments, connected by rate constants for chiral inversion was fitted to the concentration data, whereas the sedative effects were correlated with the blood concentrations of *R*- and *S*-thalidomide by means of logistic regression. PK modeling predicted that varying the infusion time of a fixed dose of *S*-thalidomide between 10 min and 6 h would have little influence on the maximal blood concentration of formed *R*-thalidomide. These effects may chiefly be exerted by *S*-thalidomide, but the enantiomers are interconverted *in vivo*

(*Continued*)

TABLE 18.3 (*Continued*)

Drug Reference	Findings
Ivabradine[67]	Ivabradine has an active metabolite S-18982. Both ivabradine and S-18982 possess bradycardic activity, although the extent of the activity of both could not be determined from the data available. A multiple ligand PD model provided the best fit to the data. The model was assessed in terms of its posterior predictive performance and was able to describe the original data adequately when used for simulation purposes
Lorazepam[68]	The PK-PD modeling of the psychomotor and mnesic effects of a single 2 mg oral dose of lorazepam in healthy volunteers showed that the parameter values derived from PK/PD modeling, and especially the EC_{50} values, may provide sensitive indices that can be used, rather than the raw data derived from PD measurements, to compare CNS effects of benzodiazepines
Cetrizine[69]	The PK-PD modeling characterized the time course of cetirizine effect on histamine-induced skin reactions (wheal and flare). The histamine-induced flare was at least 80% inhibited at the start of the second administration. Thereafter, on successive administrations, the inhibition was even more pronounced and the response control was nearly total
Isepamicin[70]	Population PK and Bayesian estimates of individual PK parameters were used to calculate various surrogate markers of isepamicin exposure to be tentatively correlated with clinical outcome and nephrotoxicity. No correlation was found between peak, AUC, or their ratio with MIC and clinical efficacy suggesting the drug should not be monitored using PK parameters
Formoterol[71]	Interactions of formoterol and theophylline were evaluated with the use of PK/PD modeling. The E_{max} values for the eosinopenic effects were fixed at 0. Effects of both drugs combined were described with a noncompetitive interaction model
Furesimide[72]	The time of maximum response to frusimide was better described using an indirect-response model than an effect-compartment model. Studying the time of maximum response after administration of different single doses of a drug may be used as a selective tool during PK/PD modeling
Piperacillin-tazobactam[73]	*In vitro* anti-infective effect of piperacillin-tazobactam (PIP-TZB) combinations on *Escherichia coli* showed through a PK/PD model that for these combinations, three-times-a-day administration is as effective as four times a day. Pharmacodynamic activity of the combinations can be prolonged by sufficiently high inhibitor concentrations
Diltiazem[74]	Functional relationship between PK and PD parameters is described by hysteresis loops with a clockwise rotation. This cannot be explained in the classical way by the time lag between central and effect compartments. The model of down regulation/tolerance development is proposed as a result supported by the finding that the shape of the hysteresis is dependent on the absorption rate of diltiazem, calculated as mean input time. Acute tolerance to dilitazem develops at least with the electrophysiological action of diltiazem after oral application and that the extent of tolerance development increases when decreasing its absorption rate. Bioequivalence assessment of diltiazem is possible using PD parameters; however, because PK/PD relationships are influenced by the absorption rate, extent parameters may be misinterpreted when rate parameters of the test formulations are different
Oxprenolol[75]	The relationship between the plasma concentration of oxprenolol and its haemodynamic effects during physical exercise was studied using the fitted PK curve; the drug effect was related to its plasma concentration using the E_{max} model
Trimazosin[76,77]	The PK and PD of trimazosin are described following both intravenous and oral administration. The IV and oral drug and metabolite (1-hydroxytrimazosin) concentration data are fitted simultaneously to the same PK model. The PD response, change in systolic blood pressure. It appears that the reduction in blood

(*Continued*)

TABLE 18.3 (*Continued*)

Drug Reference	Findings
	pressure following administration of trimazosin at steady state may be governed by the concentration of metabolite. The PK and PD profiles of intravenous trimazosin were analyzed empirically by integrated modeling techniques. The placebo-corrected fall in systolic blood pressure after 5 min standing was most appropriately characterized in all subjects by a model that incorporated effect compartments for both parent drug and its metabolite

of drug elimination from the body, and thus it is possible to see a persistent drug effect without any detectable drug concentration. The maximum response in indirect model drugs is linearly proportional to logarithm of dose unlike the direct action models where the response E_{max} is a dose-independent parameter[54] as shown in effect-compartment model. Studies on the shape of concentration–effect response in drugs that act by stimulation shows that the rate of drug delivery is crucial and thus sustained or targeted delivery of these drugs would maximize efficacy and reduce the side effects.

The indirect response model has components of physiological parameters as affected by disease, concomitant drug administration, demographic, genotypic, and other patient variables and thus provides a much better clinical evaluation opportunity; however, establishing indirect response models requires study of several doses that must be fitted simultaneously. Using a classic model and observing that EC_{50} or E_{max} are dose-dependent should provide an indication to switch to indirect response model.[55]

18.10 Examples

A large volume of published data provides examples of application of PK/PD. Table 18.3 lists some of the most recent attempts to project PK and PD in combined studies.

18.11 Conclusions and Recommendations

The integration of PK and PD principles into the drug development process to make it more rational and efficient is highly laudable. This integration relies extensively on PK/PD models to describe the relationships among dose, concentration (and, more generally, exposure), and the responses, such as surrogate markers, efficacy measures, and adverse events. Well-documented empirical and physiological PK/PD models are, increasingly, becoming available.

Population PK/PD models, which in addition to the characterization of PK and PD, involve relationships between covariates (for instance, patient characteristics such as age, body weight) and PK/PD parameters, allow us to assess and to quantify potential sources of variability in exposure and response in specific target population, even under erratic and limited sampling conditions. Often implications of significant covariate effects can be evaluated by computer simulations using the population PK/PD model.

Stochastic simulation is widely used as a tool for evaluating statistical methodology, including, for example, the evaluation of performance measures for bioequivalence assessment. Recent trends indicate expanded use of simulations to support drug development for predicting the outcome of planned clinical trials. The methodological basis for this approach is provided by population PK/PD models and random sampling techniques. Models for disease progression and behavioral features such as compliance, drop-out rates, adverse event-dependent dose reductions, and the like, have to be added to population PK/PD models to simulate real situations. Use of computer simulation helps to evaluate the consequences of design features on the safety and efficacy assessment of a drug, enabling scientists and regulators to identify statistically valid and practical study designs.

References

1. Food and Drug Administration, *Development of New Stereoisomeric Drugs*, United States Food and Drug Administration, Washington, DC, 1992, Available at www.fda.gov/cder/guidance/stereo.htm
2. Temple, R.J., A regulatory authority's opinion about surrogate endpoints, in *Clinical Measurement in Drug Evaluation*, Nimmo, W.S. and Tucker, G.T., Eds., Wiley, Indianapolis, 1995.
3. Lesko, L.J. and Atkinson, A.J., Jr., Biomarkers and surrogate endpoints — Use in drug development and regulatory decision making: criteria, validation, strategies, *Ann. Rev. Pharmacol. Toxicol.*, 41, 347–366, 2001.
4. Toutain, P.L., Pharmacokinetic/pharmacodynamic integration in drug development and dosage-regimen optimization for veterinary medicine, *AAPS PharmSci.*, 4, 2002.
5. Gieschke, R. and Steimer, J.L., Pharmacometrics: modelling and simulation tools to improve decision making in clinical drug development, *Eur. J. Drug Metab. Pharmacokinet.*, 25, 49–58, 2000.
6. Wakefield, J.C. and Bennett, J.E., The Bayesian modeling of covariates for population pharmacokinetic models, *J. Am. Statist. Ass.*, 91, 917–927, 1996.
7. Machado, S.G., Miller, R., and Hu, C., A regulatory perspective on pharmacokinetic/pharmacodynamic modeling, *Stat. Methods Med. Res.*, 8, 217–245, 1999.
8. Sheiner, L.B. and Steimer, J.L., Pharmacokinetic/pharmacodynamic modeling in drug development, *Ann. Rev. Pharmacol. Toxicol.*, 40, 67–95, 2000.
9. Aarons, L., Population pharmacokinetics: theory and practice, *Br. J. Clin. Pharmacol.*, 32, 669–670, 1991.
10. Peck, C.C., Barr, W.H., Benet, L.Z., Collins, J., Desjardins, R.E., Furst, D.E., Harter, J.G., Levy, G., Ludden, T., and Rodman, J.H., Opportunities for integration of pharmacokinetics, pharmacodynamics, and toxicokinetics in rational drug development, *Int. J. Pharm.*, 82, 9–19, 1992.
11. Niazi, S., Volume of distribution and tissue level errors in instantaneous intravenous input assumptions, *J. Pharm. Sci.*, 65, 1539–1540, 1976.
12. Teorell, T., Kinetics of distribution of substances administered to the body. I: The extravascular modes of administration, *Arch. Intern. Pharmacodyn.*, 57, 205–225, 1937.
13. Teorell, T., Kinetics of distribution of substances administered to the body. II: The intravascular mode of administration, *Arch. Intern. Pharmacodyn.*, 57, 226–240, 1937.
14. Bischoff, K.B. and Brown, R.G., Drug distribution in mammals, *Chem. Eng. Prog. Symp.*, 62, 33–45, 1966.
15. Levitt, D.G., PKQuest: volatile solutes — Application to enflurane, nitrous oxide, halothane, methoxyflurane and toluene pharmacokinetics, *BMC Anesthesiol.*, 2, 5, 2002.
16. Levitt, D.G., PKQuest: measurement of intestinal absorption and first pass metabolism — Application to human ethanol pharmacokinetics, *BMC Clin. Pharmacol.*, 2, 4, 2002.
17. Levitt, D.G., PKQuest: a general physiologically based pharmacokinetic model — introduction and application to propranolol, *BMC Clin. Pharmacol.*, 2, 5, 2002.

18. Levitt, D.G., PKQuest: capillary permeability limitation and plasma protein binding — application to human inulin, dicloxacillin and ceftriaxone pharmacokinetics, *BMC Clin. Pharmacol.,* 2, 7, 2002.
19. Cai, W. and Shao, X., A fast annealing evolutionary algorithm for global optimization, *J. Comput. Chem.,* 23, 427–435, 2002.
20. Department of Mathematical Sciences, *Maple System*, Villanova University, Villanova, PA, 2003, Available at www.math.villanova.edu/archives/maple/misc
21. Levenberg, K.A., Method for the solution of certain problems in least squares, *Q. Appl. Math.,* 2, 164–168, 1944.
22. Marquardt, D., An algorithm for least-squares estimation of non-linear parameters, *SIAM J. Appl. Math.,* 11, 431–441, 1963.
23. Levitt, D.G., The use of a physiologically based pharmacokinetic model to evaluate deconvolution measurements of systemic absorption, *BMC Clin. Pharmacol.*, 3, 1, 2003.
24. De Nicolao, G., Sparacino, G., and Cobelli, C., Nonparametric input estimation in physiological systems, problems, methods, and case studies, *Automatica,* 33, 851–870, 1997.
25. Debord, J., Risco, E., Harel, M., Le Meur, Y., Buchler, M., Lachatre, G., Le Guellec, C., and Marquet, P., Application of a gamma model of absorption to oral cyclosporin, *Clin. Pharmacokinet.,* 40, 375–382, 2001.
26. Veng-Pedersen, P., Linear and nonlinear system approaches in pharmacokinetics: how much do they have to offer? I. General considerations, *J. Pharmacokinet. Biopharm.,* 16, 413–472, 1988.
27. Akaike, H., A new look at the statistical model identification, *IEEE Trans. Automat. Contr.,* 19, 716–723, 1974.
28. Craven, P. and Wahba, G., Smoothing noisy data with spline functions, *Num. Math.,* 31, 377–403, 1979.
29. Gillespie, W.R. and Veng-Pedersen, P., A polyexponential deconvolution method: evaluation of the gastrointestinal bioavailability and mean in vivo dissolution time of some ibuprofen dosage forms, *J. Pharmacokinet. Biopharm.,* 13, 289–307, 1985.
30. Verotta, D., Concepts, properties, and applications of linear systems to describe distribution, identify input, and control endogenous substances and drugs in biological systems, *Crit. Rev. Biomed. Eng.,* 24, 73–139, 1996.
31. Verotta, D., Estimation and model selection in constrained deconvolution, *Ann. Biomed. Eng.,* 21, 605–620, 1993.
32. Sparacino, G. and Cobelli, C., Deconvolution of physiological and pharmacokinetic data: comparison of algorithms on benchmark problems, in *Modeling and Control in Biomedical Systems,* Linkens, D.A. and Carson, E., Eds., Elsevier, Oxford, 1997, pp. 151–153.
33. Figg, W.D., Dukes, G.E., Lesesne, H.R., Carson, S.W., Songer, S.S., Pritchard, J.F., Hermann, D.J., Powell, J.R., and Hak, L.J., Comparison of quantitative methods to assess hepatic function: Pugh's classification, indocyanine green, antipyrine, and dextromethorphan, *Pharmacotherapy,* 15, 693–700, 1995.
34. Testa, R., Caglieris, S., Risso, D., Arzani, L., Campo, N., Alvarez, S., Giannini, E., Lantieri, P.B., and Celle, G., Monoethylglycinexylidide formation measurement as a hepatic function test to assess severity of chronic liver disease, *Am. J. Gastroenterol.,* 92, 2268–2273, 1997.
35. Tang, H.-S. and Hu, O.Y.-P., Assessment of liver function using a novel galactose single point method, *Digestion,* 52, 222–231, 1992.
36. Zakim, D. and Boyer, T.D., *Hepatology: A Textbook of Liver Disease*, W.B. Saunders Co., Philadelphia, PA, 1996.
37. Pugh, R.N., Murray-Lyon, I.M., Dawson, J.L., Pietroni, M.C., and Williams, R., Transection of the oesophagus for bleeding oesophageal varices, *Br. J. Surg.,* 60, 646–649, 1973.
38. Wiesner, R.H., Grambsch, P.M., Dickson, E.R., Ludwig, J., MacCarty, R.L., Hunter, E.B., Fleming, T.R., Fisher, L.D., Beaver, S.J., and LaRusso, N.F., Primary sclerosing cholangitis: natural history, prognostic factors and survival analysis, *Hepatology,* 10, 430–436, 1989.
39. Carithers, R.L., Jr., Herlong, H.F., Diehl, A.M., Shaw, E.W., Combes, B., Fallon, H.J., and Maddrey, W.C., Methylprednisolone therapy in patients with severe alcoholic hepatitis, *Ann. Intern. Med.,* 110, 685–690, 1989.
40. Maddrey, W.C., Boitnott, J.K., Bedine, M.S., Weber, F.L., Jr., Mezey, E., and White, R.I., Jr., Corticosteroid therapy of alcoholic hepatitis, *Gastroenterology,* 75, 193–199, 1978.

41. Toutain, P.L., Pharmacokinetic/pharmacodynamic integration in drug development and dosage-regimen optimization for veterinary medicine, *AAPS PharmSci.,* 4, 1–29, 2002.
42. Holford, N.H. and Sheiner, L.B., Pharmacokinetic and pharmacodynamic modeling in vivo, *Crit. Rev. Bioeng*., 5, 273–322, 1981.
43. Holford, N.H. and Sheiner, L.B., Kinetics of pharmacologic response, *Pharmacol Ther.,* 16, 143–166, 1982.
44. Hammarlund-Udenaes, M. and Benet, L.Z., Furosemide pharmacokinetics and pharmacodynamics in health and disease: an update, *J. Pharmacokinet. Biopharm*., 17, 1–46, 1989.
45. Landoni, M.F. and Lees, P., Pharmacokinetic/pharmacodynamic modeling of non-steroidal anti-inflammatory drugs, *J. Vet. Pharmacol. Ther*., 20, 118–120, 1997.
46. Karlsson, M.O., Molnar, V., Bergh, J., Freijs, A., and Larsson, R., A general model for time–dissociated pharmacokinetic–pharmacodynamic relationships exemplified by paclitaxel myelosuppression, *Clin. Pharmacol. Ther.,* 63, 11–25, 1998.
47. Campbell, D.B., The use of kinetic-dynamic interactions in the evaluation of drugs, *Psychopharmacology,* 100, 433–450, 1990.
48. Mattie, H., Antibiotic efficacy in vivo predicted by in vitro activity, *Int. J. Antimicrob. Agents,* 14, 91–98, 2000.
49. Mager, D.E. and Jusko, W.J., Pharmacodynamic modeling of time-dependent transduction systems, *Clin. Pharmacol. Ther*., 70, 210–216, 2001.
50. Fuseau, E. and Sheiner, L.B., Simultaneous modeling of pharmacokinetics and pharmacodynamics with a nonparametric pharmacodynamic model, *Clin. Pharmacol. Ther.,* 35, 733–741, 1984.
51. Sheiner, L.B. and Beal, S.L., Bayesian individualisation of pharmacokinetics: simple implementation and comparison with non-Bayesian methods, *J. Pharm. Sci.,* 71, 1344–1348, 1982.
52. Jusko, W.J. and Ko, H.C., Physiologic indirect response models characterize diverse types of pharmacodynamic effects, *Clin. Pharmacol. Ther*., 56, 406–419, 1994.
53. Sharma, A. and Jusko, W.J., Characterization of four basic models of indirect pharmacodynamic responses, *J. Pharmacokinet. Biopharm*., 24, 611–635, 1996.
54. Gobburu, J.V. and Jusko, W.J., Role of dosage regimen in controlling indirect pharmacodynamic responses, *Adv. Drug Deliv. Rev*., 46, 45–57, 2001.
55. Levy, G., Mechanism-based pharmacodynamic modeling, *Clin. Pharmacol. Ther*., 56, 356–358, 1994.
56. Aymard, G., Berlin, I., de Brettes, B., and Diquet, B., Pharmacokinetic–pharmacodynamic study of apomorphine's effect on growth hormone secretion in healthy subjects, *Fundam. Clin. Pharmacol*., 17, 473–481, 2003.
57. Varlet-Marie, E., Gaudard, A., Audran, M., Gomeni, R., and Bressolle, F., Pharmacokinetic–pharmacodynamic modeling of recombinant human erythropoietin in athletes, *Int. J. Sports Med.,* 24, 252–257, 2003.
58. Hornestam, B., Jerling, M., Karlsson, M.O., and Held, P.; DAAf Trial Group, Intravenously administered digoxin in patients with acute atrial fibrillation: a population pharmacokinetic/pharmacodynamic analysis based on the digitalis in acute atrial fibrillation trial, *Eur. J. Clin. Pharmacol.,* 58, 747–755, 2003.
59. Fujita, T., Matsumoto, Y., Kimura, T., Yokota, S., Sawada, M., Majima, M., Ohtani, Y., and Kumagai, Y., Pharmacokinetics and pharmacodynamics of TF-505, a novel nonsteroidal 5α-reductase inhibitor, in normal subjects treated with single or multiple doses, *Br. J. Clin. Pharmacol*., 54, 283–294, 2002.
60. Svensson, U.S., Alin, H., Karlsson, M.O., Bergqvist, Y., and Ashton, M., Population pharmacokinetic and pharmacodynamic modeling of artemisinin and mefloquine enantiomers in patients with falciparum malaria, *Eur. J. Clin. Pharmacol*., 58, 339–351, 2002.
61. Nicolau, D.P., Pharmacodynamic rationale for short-duration antibacterial therapy, *J. Infect*., 44, 17–23, 2002.
62. Lin, S. and Chien, Y.W., Pharmacokinetic–pharmacodynamic modeling of insulin: comparison of indirect pharmacodynamic response with effect-compartment link models, *J. Pharm. Pharmacol.,* 54, 791–800, 2002.

63. Csajka, C., Buclin, T., Fattinger, K., Brunner, H.R., and Biollaz, J., Population pharmacokinetic–pharmacodynamic modelling of angiotensin receptor blockade in healthy volunteers, *Clin. Pharmacokinet.,* 41, 137–152, 2002.
64. Miyazaki, M., Mukai, H., Iwanaga, K., Morimoto, K., and Kakemi, M., Pharmacokinetic–pharmacodynamic modeling of human insulin: validity of pharmacological availability as a substitute for extent of bioavailability, *J. Pharm. Pharmacol*., 53, 1235–1246, 2001.
65. Zhou, H., Chen., T.L., Marino, M., Lau, H., Miller, T., Kalafsky, G., and McLeod, J.F., Population PK and PK/PD modeling of microencapsulated octreotide acetate in healthy subjects, *Br. J. Clin. Pharmacol*., 50, 543–552, 2000.
66. Eriksson, T., Bjorkman, S., Roth, B., and Hoglund, P., Intravenous formulations of the enantiomers of thalidomide: pharmacokinetic and initial pharmacodynamic characterization in man, *J. Pharm. Pharmacol.,* 52, 807–817, 2000.
67. Duffull, S.B. and Aarons, L., Development of a sequential linked pharmacokinetic and pharmacodynamic simulation model for ivabradine in healthy volunteers, *Eur. J. Pharm. Sci.,* 10, 275–284, 2000.
68. Blin, O., Jacquet, A., Callamand, S., Jouve, E., Habib, M., Gayraud, D., Durand, A., Bruguerolle, B., and Pisano, P., Pharmacokinetic–pharmacodynamic analysis of mnesic effects of lorazepam in healthy volunteers, *Br. J. Clin. Pharmacol*., 48, 510–512, 1999.
69. Urien, S., Tillement, J.P., Ganem, B., and Kuch, M.D., A pharmacokinetic–pharmacodynamic modeling of the antihistaminic (H1) effects of cetirizine, *Int. J. Clin. Pharmacol. Ther*., 37, 499–502, 1999.
70. Tod, M., Minozzi, C., Beaucaire, G., Ponsonnet, D., Cougnard, J., and Petitjean, O., Isepamicin in intensive care unit patients with nosocomial pneumonia: population pharmacokinetic–pharmacodynamic study, *J. Antimicrob. Chemother*., 44, 99–108, 1999.
71. van den Berg, B.T., Derks, M.G., Koolen, M.G., Braat, M.C., Butter, J.J., and van Boxtel, C.J., Pharmacokinetic/pharmacodynamic modeling of the eosinopenic and hypokalemic effects of formoterol and theophylline combination in healthy men, *Pulm. Pharmacol. Ther*., 12, 185–192, 1999.
72. Wakelkamp, M., Alvan, G., and Paintaud, G., The time of maximum effect for model selection in pharmacokinetic–pharmacodynamic analysis applied to frusemide, *Br. J. Clin. Pharmacol*., 45, 63–70, 1998.
73. Dalla Costa, T., Nolting, A., Rand, K., and Derendorf, H., Pharmacokinetic–pharmacodynamic modeling of the *in vitro* antiinfective effect of piperacillin-tazobactam combinations, *Int. J. Clin. Pharmacol. Ther.,* 35, 426–433, 1997.
74. Luckow, V. and Della Paschoa, O., PK/PD modeling of high-dose diltiazem—absorption-rate dependency of the hysteresis loop, *Int. J. Clin. Pharmacol. Ther*., 35, 418–425, 1997.
75. Koopmans, R., Oosterhuis, B., Karemaker, J.M., Wemer, J., and van Boxtel, C.J., Pharmacokinetic–pharmacodynamic modeling of oxprenolol in man using continuous noninvasive blood pressure monitoring, *Eur. J. Clin. Pharmacol*., 34, 395–400, 1988.
76. Kelman, A.W., Meredith, P.A., Elliott, H.L., and Reid, J.L., Modeling the pharmacokinetics and pharmacodynamics of trimazosin, *Biopharm. Drug Dispos*., 7, 373–388, 1986.
77. Meredith, P.A., Kelman, A.W., Elliott, H.L., and Reid, J.L., Pharmacokinetic and pharmacodynamic modeling of trimazosin and its major metabolite, *J. Pharmacokinet. Biopharm*., 11, 323–325, 1983.

19

Pharmaceutics and Compounding Issues in New Drug Development and Marketing

Loyd V. Allen

CONTENTS

19.1 Introduction

Pharmaceutics-based pharmaceutical compounding is an integral part of providing pharmaceuticals and is essential to the provision of contemporary health care.[1] Pharmaceutical compounding has an interesting relationship in the development of new drugs and is supported by the pharmaceutical sciences. Compounding is a professional prerogative that pharmacists have performed since the beginning of the profession. Compounding can be as simple as the addition of a liquid to a manufactured drug powder, or as complex as the preparation of a multicomponent parenteral nutrition solution or a multicomponent transdermal gel. In general, compounding differs from manufacturing in that compounding involves a specific practitioner–patient–pharmacist relationship, the preparation of a relatively small quantity of medication, and different conditions of sale (i.e., specific prescription orders). The pharmacist is responsible for compounding preparations of acceptable strength, quality, and purity with appropriate packaging and labeling in accordance with good pharmacy practices, official standards, and current scientific principles (pharmaceutics).

19.2 Definitions

Pharmaceutical compounding is defined as the act of preparing, mixing, assembling, packaging, and labeling of a drug or device as a result of a practitioner's prescription drug order, or the initiative based on the practitioner–patient–pharmacist relationship in the course of professional practice, or for the purpose of, or as an incident to, research, teaching, or chemical analysis and not for sale or dispensing. Compounding also includes the preparation of drugs or devices in anticipation of prescription drug orders based on routine, regularly observed prescribing patterns.[2–4]

Pharmaceutical manufacturing is defined as the production, preparation, propagation, conversion, and processing of a drug or device, either directly or indirectly, by extraction from substances of natural origin or independently by means of chemical or biological synthesis. This includes any packaging or repackaging of the substance(s), labeling or relabeling of its

container, and the promotion and marketing of such drugs or devices. Manufacturing also includes the preparation and promotion of commercially available products from bulk compounds for resale by pharmacies, practitioners, or other people.[2,3]

19.3 Relationship between Compounding and Manufacturing

The 5000 year history and heritage of pharmacy has centered around the provision of pharmaceutical products for patients. Pharmacists are the only health care professionals who possess the knowledge and skill required for compounding and preparing medications to meet the unique and individual needs of patients. The responsibility of extemporaneously compounding safe, effective prescription products for patients who require special care has always been fundamental to the pharmacy profession.

The 19th century did not see an end to compounding but was impacted by new technology. In the 20th century, a broad knowledge of compounding was still essential for about 80% of the prescriptions dispensed in the 1920s. Although pharmacists increasingly relied on commercially available products for dispensing, and chemicals purchased from the manufacturers to make up prescriptions, there still remained much to be done *secundum artem*.[5]

The pharmaceutical industry began to assume the production of most medications and in many ways this has provided superior service and utilized new methods and a vast array of innovative products that could not have been provided on a one-on-one basis. Research and development have historically been the hallmarks of pharmaceutical manufacturers. However, the very nature of providing millions of doses of a product requires that the dosage forms (such as capsules, tablets, and suppositories) along with the dose amounts (individual strengths of each dose) be limited and result in a unilateral approach to therapy. Today, it is simply not economical for a pharmaceutical company to produce a product in numerous different doses or dosage forms to meet the needs of the entire range of individuals receiving therapy. Products are designed that will meet the majority of patient needs, but the very nature of the process cannot meet all patient needs.

Pharmaceutical compounding has recently been increasing for a number of reasons, including:

- Availability of a limited number of dosage forms for most drugs
- A limited number of strengths of most drugs, home health care, and hospice
- The nonavailability of drug products or combinations, discontinued drugs, drug shortages, and orphan drugs
- New therapeutic approaches and special patient populations (such as pediatrics, geriatrics, bioidentical hormone replacement therapy for postmenopausal women, pain management, and dental patients)
- Environmentally and cosmetic sensitive patients, sports injuries, and veterinary compounding, including small, large, herd, exotic, and companion animals

Over the years it has been interesting to note that many compounded products eventually become commercially available products. Recent examples might include fentanyl lozenges, minoxidil topical solution, nystatin lozenges, clindamycin topical solution, tetracaine–adrenalin–cocaine (TAC) solution, dihydroergotamine mesylate nasal spray, buprenorphine nasal spray, buffered hypertonic saline solution, and erythromycin topical solution as well as numerous other dermatological and pediatric oral liquids and some

premixed intravenous solutions. It is inevitable that a product will be manufactured when it becomes economically profitable for a pharmaceutical manufacturer to produce it. Likewise, when a commercially available product is no longer profitable for a manufacturer, they tend to drop it and pharmacists begin compounding it again.

19.4 Pharmaceutics

Drug substances are not generally given by themselves, but rather as part of a prepared formulation in combination with excipients that serve varied and specialized pharmaceutical functions. Through selective use of these excipients, various types of dosage forms result. The excipients are used to solubilize, suspend, thicken, dilute, emulsify, stabilize, preserve, color, flavor, and fashion medicinal agents into efficacious and appealing dosage forms. Each type of dosage form is unique in its physical and pharmaceutical characteristics. These preparations provide the manufacturing and compounding pharmacist with the challenges of formulation and the physician with the choice of drug and drug delivery system to prescribe. The general area of study concerned with the formulation, manufacture, stability, and effectiveness of pharmaceutical dosage forms is termed pharmaceutics.

The proper design and formulation of a dosage form requires considering the physical, chemical, and biological characteristics of all the drug substances and excipients that are used in fabricating the product. All the ingredients utilized must be compatible to produce a drug product that is stable, efficacious, attractive, easy to administer, and safe. The formulation should be prepared under appropriate measures of quality control and packaged in containers that maintain product stability. The product should be properly labeled and stored under conditions that contribute to maximum shelf-life.

19.5 The Need for Dosage Forms

The potent nature and low dosage of most of the drugs in use today require a dosage form for a number of reasons, including:

- The protection of a drug substance from the destructive influences of atmospheric oxygen or humidity (such as coated tablets, sealed ampuls)
- The protection of a drug substance from the destructive influence of gastric acid after oral administration (such as enteric-coated tablets)
- The concealment of the bitter, salty, or offensive taste or odor of a drug substance (such as capsules, coated tablets, flavored syrups)
- The provision of liquid preparations of substances that are either insoluble or unstable in the desired vehicle (such as suspensions)
- The provision of clear liquid dosage forms of substances (such as syrups, solutions)
- The provision of rate-controlled drug action (such as various controlled-release tablets, capsules, and suspensions)
- The provision of optimal drug action from topical administration sites (such as ointments, creams, transdermal patches, ophthalmic, ear, and nasal preparations)
- The provision of the insertion of a drug into one of the body's orifices (such as rectal or vaginal suppositories)

- The provision for the placement of drugs directly into the bloodstream or into body tissues (such as injections)
- The provision of optimal drug action through inhalation therapy (such as inhalants and inhalation aerosols)

19.6 General Considerations in Dosage-Form Design

Prior to formulating a drug substance into a dosage form, the desired product type must be detemined for planning the product formulation activities. Then, various initial formulations are developed and then evaluated for selected parameters, such as drug-release profile, bioavailability, clinical effectiveness, and for any scale-up problems. The best formulation is selected and becomes the master formula. Each batch of the product subsequently prepared must meet the specifications established in this master formula.

A wide variety of dosage forms is available in which a medicinal agent may be placed for convenient and efficacious treatment of disease. Optimally, a pharmaceutical manufacturer would prepare a drug substance in several dosage forms and strengths for different patient needs. Before a medicinal agent is formulated into one or more dosage forms, several factors must be considered, including therapeutic matters such as, the nature of the illness, the manner in which it is treated (locally or systemically), and the age and anticipated condition of the patient.

If the medication is intended for systemic use and oral administration is desired, tablets and capsules are usually prepared. These dosage units are easily handled by the patient and are most convenient in the self-administration of medication. If a drug substance has an application in an emergency situation in which the patient may be comatose or unable to take oral medication, an injectable form of the medication may also be prepared. Examples of therapeutic situations affecting dosage-form design include the preparation of agents for motion sickness, nausea, and vomiting, into tablets and skin patches for prevention and suppositories and injections for treatment.

Patient age also plays a role in dosage-form design. For infants and children, younger than 5 years of age, pharmaceutical liquids rather than solid dosage forms are preferred for oral administration. These liquids can be flavored aqueous solutions, syrups or suspensions, and are usually administered directly into the infant's or child's mouth by drop, spoon, or oral dispenser or incorporated into the child's food. A single liquid pediatric preparation may be used for infants and children of all ages, with the dose of the drug varied by the volume administered. Sometimes two different strengths are available, one for infants and one for children.

Some children, and even adults, may have difficulty swallowing solid dosage forms, especially uncoated tablets. For this reason, some medications are formulated as chewable tablets that can be broken up in the mouth before swallowing. Many of these tablets are comparable in texture to an after-dinner mint and break down into a pleasant tasting, creamy material. New, rapidly disintegrating and dissolving tablets are available that dissolve in the mouth in just a few seconds, allowing the patient to take a tablet but actually swallow a liquid. Capsules have been found by many to be more easily swallowed than whole tablets. If a capsule is allowed to become moist in the mouth before swallowing, it becomes slippery and slides down the throat more readily with a glass of water. Medications intended for the elderly are commonly formulated into oral liquids or may be extemporaneously prepared into an oral liquid by the pharmacist.

Since many patients, particularly the elderly, take multiple medications daily, proper identification of the medication is made easier by the medication having a distinctive size, shape, and color of solid dosage forms. Frequent errors in taking medications among the elderly occur because of their multiple-drug therapy and reduced eyesight. Dosage forms that allow reduced frequency of administration without sacrifice of efficiency are particularly advantageous.

During the process of formulating a drug substance into a proper dosage form, research pharmacists employ knowledge that has been gained through experience with other chemically similar drugs and through the proper utilization of the disciplines of physical, chemical, biological, and pharmaceutical sciences. Preliminary steps for any new formulation involves studies to collect basic information on the physical and chemical characteristics of the drug substance to be prepared into pharmaceutical dosage forms; this is called preformulation and is needed before actual product formulation begins.

19.6.1 Preformulation Studies

A drug substance must be chemically and physically characterized prior to incorporation into a new dosage form. Preformulation work provides the type of information needed to define the nature of the drug substance, as thoroughly as possible, and this then provides the framework for the drug's combination with pharmaceutic ingredients in the fabrication of a dosage form.

19.6.2 Physical Description

One begins with an understanding of the physical description of a drug substance. The majority of drug substances in use today occur as solid materials, mostly white powders, with most of them being pure chemical compounds of either crystalline or amorphous constitution. The purity of the chemical substance is essential for its identification as well as for the evaluation of its chemical, physical, and biological properties. Chemical properties include structure, form, and reactivity. Physical properties include characteristics such as its physical description, particle size, crystalline structure, melting point, and solubility. Biological properties relate to its ability to get to a site of action and elicit a biologic response.

Drugs can be beneficially used in therapy as solids, liquids, and gases. The form most commonly used in drug therapy is solid, followed by liquid and then gas. Liquid drugs can pose an interesting challenge in dosage-form development since many of the liquids are volatile substances and, as such, must be physically sealed from the atmosphere to prevent their loss.

Liquid drugs pose interesting problems when required to be formulated into tablet form, the most popular form of oral medication, without undertaking chemical modification of the drug. An exception to this is the liquid drug nitroglycerin, which is formulated into sublingual tablets that disintegrate within seconds after placing under the tongue.In most cases, when a liquid drug is to be administered orally and a solid dosage form is desired, two approaches are used. First, the liquid substance may be sealed in a soft gelatin capsule or, secondly, the liquid drug may be developed into a solid ester or salt form that will be suitable for tableting or drug encapsulating. Another unique approach to formulate liquids into solids is by mixing the drug with a solid or a melted semisolid material, such as a high-molecular-weight polyethylene glycol. The melted mixture is poured into hard gelatin capsule where it will harden, and the capsule sealed.

Formulation and stability difficulties occur less frequently with solid dosage forms than with liquid pharmaceutical preparations. For this reason, most new drugs are marketed as

tablets or dry-filled capsules. Later, when the pharmaceutical problems are resolved, a liquid form of the same drug may be marketed. This procedure, when practiced, is doubly advantageous, because for the most part physicians and patients alike prefer small, generally tasteless, accurately dosed tablets or capsules to the analogous liquid forms. Tablets and capsules comprise the most common dosage form dispensed by pharmacists, with tablets dispensed twice as frequently as capsules.

19.6.3 Microscopic Examination

Microscopic examination of the raw drug substance provides an indication of particle size and particle size range of the raw material as well as the crystal structure. Photomicrographs of the initial and subsequent batch lots of the drug substance can provide important information should problems arise in formulation processing. These problems are attributable to changes in the particle or crystal characteristics of the drug. Particle flow can be important as spherical- and oval-shaped powders flow more easily than needle-shaped powders and make processing easier.

19.6.4 Melting Point Depression

A defined melting point or melting range is a characteristic of a pure substance. If not pure, the substance will exhibit a depressed melting point. This phenomenon is commonly used to determine the purity of a drug substance and, in some cases, the compatibility of various substances before inclusion in the same dosage form.

19.6.5 Phase Diagrams

Phase diagrams are often constructed to provide a visual picture of the existence and extent of the presence of solid and liquid phases in binary, ternary, and other mixtures of substances. Phase diagrams are normally two-component (binary) representations but multicomponent phase diagrams can also be constructed. Interactions between active substances and excipients can often be evaluated using phase diagrams.

19.6.6 Particle Size

Particle size can affect certain physical and chemical properties of drug substances, including drug dissolution rate, bioavailability, content uniformity, taste, texture, color, and stability. Also, properties such as flow characteristics and sedimentation rates, among others, are important factors related to particle size. The effect of particle size on formulation and product efficacy must be determined as early as possible. Of special interest is the effect of particle size on the drug's absorption. Particle size significantly influences the oral absorption profiles of certain drugs. Also, content uniformity in solid dosage forms depends to a large degree on particle size and the equal distribution of the active ingredient throughout the formulation.

19.6.7 Polymorphism

Crystalline or amorphous forms of the drug substance can affect product efficacy. Polymorphic forms usually exhibit different physical–chemical properties, including melting point and solubility. The occurrence of polymorphic forms with drugs is relatively

common and it has been estimated that polymorphism is exhibited by at least one third of all organic compounds.

Amorphous or noncrystalline forms can exist and the energy required for a molecule of drug to escape from a crystal is much greater than that required to escape from an amorphous powder. Therefore, the amorphous form of a compound is always more soluble than a corresponding crystal form.

Crystalline structure, polymorphism, and solvate form are important preformulation considerations. The changes in crystal characteristics can influence bioavailability, chemical, and physical stability, and have important implications in dosage-form process functions. For example, it can be a significant factor relating to the tableting processes due to flow and compaction behaviors, among others. Techniques used in determining crystalline properties include hot-stage microscopy, thermal analysis, infrared spectroscopy, and x-ray diffraction.

19.6.8 Solubility

For a drug to enter the systemic circulation and exert a therapeutic effect, it must first be in a solution. A drug must possess some aqueous solubility for therapeutic efficacy. Relatively insoluble compounds often exhibit incomplete or erratic absorption. If the solubility of the drug substance is low, attention may be required to improve its solubility. Methods to accomplish this will depend on the chemical nature of the drug and the type of drug product under consideration. Chemical modification of the drug into salt or ester forms is a technique frequently used to obtain more soluble compounds. Although solubility is normally considered a physicochemical constant, small increases in solubility can be accomplished by particle size reduction.

A drug's solubility is usually determined by the equilibrium solubility method, by which an excess of the drug is placed in a solvent and shaken at a constant temperature over a prolonged period of time until equilibrium is obtained. Chemical analysis of the drug content in solution is performed to determine degree of solubility.

Another technique, if the drug is to be formulated into a liquid product, involves the adjustment of the pH of the solvent in which the drug is to be dissolved to enhance solubility. However, there are many drug substances for which pH adjustment is not an effective means of improving solubility. Weak acidic or basic drugs may require extremes of pH that exceed accepted physiologic limits or may cause stability problems with formulation ingredients. Adjustment of pH usually has little effect on the solubility of nonelectrolytes. In many cases, it is desirable to utilize cosolvents or other techniques such as complexation, micronization, or solid dispersion to improve aqueous solubility.

19.6.9 Dissolution

Variations in the biological activity of a drug substance may be brought about by the rate at which it becomes available to the organism. Oftentimes, dissolution rate, or the time it takes for the drug to dissolve in the fluids at the absorption site, is the rate-limiting step in the absorption process. This is true for drugs administered orally in solid forms such as tablets, capsules, or suspensions, as well as drugs administered intramuscularly in the form of pellets or suspensions. When the dissolution rate is the rate-limiting step, anything affecting dissolution will also affect absorption. Consequently, dissolution rate can affect the onset, intensity, and duration of response, and control the overall bioavailability of the drug from the dosage form.

The dissolution rate of drugs may be improved by decreasing the drug's particle size. The most effective means of obtaining higher dissolution rates is to use a highly water-soluble

salt of the parent substance. Although a soluble salt of a weak acid will subsequently precipitate as the free acid in the bulk phase of an acidic solution, such as gastric fluid, it will do so in the form of fine particles with a large surface area.

19.6.10 Membrane Permeability

Generally, to produce a biological response, a drug molecule must first cross at least one biological membrane. The biological membrane acts as a lipid barrier to most drugs and permits the absorption of lipid-soluble substances by passive diffusion while lipid-insoluble substances can diffuse if at all across the barrier only with considerable difficulty. The interrelationship of the dissociation constant, lipid solubility, and pH at the absorption site and absorption characteristics of various drugs are the basis of the pH-partition theory.

19.6.11 Partition Coefficient

In formulation development, the octanol–water partition coefficient is commonly used; it is defined as:

P = concentration of drug in octanol/concentration of drug in water

P is dependent on the drug concentration only if the drug molecules have a tendency to associate in solution. For an ionizable drug, the following equation is applicable:

P = concentration of drug in octanol/$(1 - \alpha)$(concentration of drug in water)

where α equals the degree of ionization.

A drug substance generally must have both water- and lipid-soluble characteristics to be optimally effective.

19.6.12 Dissociation Constant or pK_a

The extent of ionization has an important effect on the formulation and pharmacokinetic parameters of the drug. The extent of dissociation or ionization is, in many cases, highly dependent on the pH of the medium containing the drug. In formulation, often the vehicle is adjusted to a certain pH to obtain a certain level of ionization of the drug for solubility and stability purposes. In the pharmacokinetic area, the extent of ionization of a drug is an important affector of its extent of absorption, distribution, and elimination. Dissociation constant or pK_a is usually determined by potentiometric titration. For a practicing pharmacist, it is important to predict precipitation in admixtures and when calculating the solubility of drugs at certain pH values.

19.6.13 Stability

Stability is defined as the extent to which a product retains, within specified limits and throughout its period of storage and use (such as its shelf-life), the same properties and characteristics that it possessed at the time of its manufacture.

There are five types of stability of concern to formulation pharmacists in both manufacturing and compounding:

- *Chemical*: Each active ingredient retains its chemical integrity and labeled potency, within the specified limits.
- *Physical*: The original physical properties, including appearance, palatability, uniformity, dissolution, and suspendability are retained.

- *Microbiological*: Sterility or resistance to microbial growth is retained according to the specified requirements. Antimicrobial agents that are present retain effectiveness within specified limits.
- *Therapeutical*: The therapeutic effect remains unchanged.
- *Toxicological*: No significant increase in toxicity occurs.

Chemical stability is important for selecting storage conditions (such as temperature, light, and humidity), selecting the proper container for dispensing (glass vs. plastic, clear vs. amber, or opaque and cap liners), and for anticipating interactions when mixing drugs and dosage forms. Stability and expiration dating are based on reaction kinetics.

One of the most important activities of preformulation work is the evaluation of the physical and chemical stability of the pure drug substance and the final dosage form. It is essential that these initial studies be conducted using drug samples of known purity. The presence of impurities can lead to erroneous conclusions in such evaluations. Stability studies conducted in the preformulation phase include solid-state stability of the drug alone, solution-phase stability, and stability in the presence of expected excipients. Initial investigation begins through knowledge of the drug's chemical structure, which allows the preformulation scientist to anticipate the possible degradation reactions.

19.6.14 Chemical Stability

In considering chemical stability of a pharmaceutical, one must evaluate the reaction order and reaction rate. The reaction order may be the overall order (the sum of the exponents of the concentration terms of the rate expression), or the order with respect to each reactant (the exponent of the individual concentration term in the rate expression). The reaction rate expression is a description of the drug concentration with respect to time. Most commonly, zero- and first-order reactions are encountered in pharmacy.

Chemical instability of medicinal agents may take many forms, because the drugs in use today are of such diverse chemical constitution. Chemically, drug substances are alcohols, phenols, aldehydes, ketones, esters, ethers, acids, salts, alkaloids, glycosides, and others, each with reactive chemical groups having different stability characteristics. Chemically, the most frequently encountered destructive processes involve hydrolysis and oxidation.

Hydrolysis is a solvolysis process in which (drug) molecules interact with water molecules to yield breakdown products of different chemical constitution. The process of hydrolysis is probably the most important single cause of drug decomposition, mainly because a great number of medicinal agents are esters or contain such other groupings as substituted amides, lactones, and lactams, which are susceptible to the hydrolytic process.[6]

The oxidative process is destructive to many pharmaceuticals, including aldehydes, alcohols, phenols, sugars, alkaloids, and unsaturated fats and oils. Many of the oxidative changes in pharmaceutical preparations have the character of autoxidations. Autoxidations occur spontaneously under the initial influence of atmospheric oxygen and proceed slowly at first and then more rapidly as the process continues.

In drug product formulation work, steps are taken to reduce or prevent the occurrence of drug substance deterioration due to hydrolysis, oxidation, and other processes.

19.6.15 Enhancing Stability of Drug Products

Many pharmaceutic ingredients may be utilized in preparing the desired dosage form of a drug substance. Some of these agents may be used to achieve the desired physical and chemical characteristics of the product or to enhance its appearance, odor, and taste. Other

substances may be used to increase the stability of the drug substance, particularly against the hydrolytic and oxidative processes. In each instance, the added pharmaceutic ingredient must be compatible with and must not detract from the stability of the drug substance in the particular dosage form prepared.

There are several approaches to the stabilization of pharmaceutical preparations containing drugs subject to deterioration by hydrolysis, the most obvious of which is the reduction or the elimination of water from the pharmaceutical system. Even solid dosage forms containing water-labile drugs must be protected from the humidity of the atmosphere. This may be accomplished by applying a waterproof protective coating over tablets or by enclosing and maintaining the drug in tightly closed containers. In liquid preparations, water can frequently be replaced or reduced in the formulation through the use of substitute liquids such as glycerin, propylene glycol, and alcohol. In certain injectable products, anhydrous vegetable oils may be used as the drug's solvent to reduce the chance of hydrolytic decomposition.

Decomposition by hydrolysis may be prevented for other drugs to be administered in liquid form by suspending them in a nonaqueous vehicle rather than by dissolving them in an aqueous solvent. Alternatively, particularly for certain unstable antibiotic drugs, when an aqueous preparation is desired, the drug may be formulated in a dry form for later reconstitution by adding a specified volume of purified water, just before dispensing or administering. Storage under refrigeration is advisable for most preparations considered unstable due to hydrolytic causes. Together with temperature, pH is a major determinant in the stability of a drug prone to hydrolytic decomposition. The hydrolysis of most drugs is dependent upon the relative concentrations of the hydroxyl and hydronium ions, and a pH at which each drug is optimally stable can be easily determined. For most hydrolyzable drugs, the pH of optimum stability is on the acid side, somewhere between pH 5 and 6. Therefore, through judicious use of buffering agents, the stability of otherwise unstable compounds can be increased.

The oxidative process can be minimized or slowed by the use of antioxidants, which react with one or more compounds in the drug to prevent progress of the chain reaction. In general, antioxidants act by providing electrons and easily available hydrogen atoms that are accepted more readily by the free radicals than those of the drug being protected. Various antioxidants are employed in pharmacy. Among these, more frequently used in aqueous preparations are sodium sulfite (Na_2SO_3, at high pH values), sodium bisulfite ($NaHSO_3$, at intermediate pH values), sodium metabisulfite ($Na_2S_2O_5$, at low pH values), hypophosphorous acid (H_3PO_2), and ascorbic acid. In oleaginous (oily or unctuous) preparations, α-tocopherol, butylhydroxyanisole, and ascorbyl palmitate find application.

The proper use of antioxidants involves their specific application only after appropriate biomedical and pharmaceutical studies. In certain instances other pharmaceutical additives can inactivate a given antioxidant when used in the same formulation. In other cases certain antioxidants can react chemically with the drugs they were intended to stabilize, without a noticeable change in the appearance of the preparation. Because the stability of oxidizable drugs may be adversely affected by oxygen, certain pharmaceuticals may require an oxygen-free atmosphere during their preparation and storage. To avoid these exposures, oxygen-sensitive drugs may be prepared in the dry state and they, as well as liquid preparations, may be packaged in sealed containers with the air replaced by an inert gas such as nitrogen. This is a common practice in the commercial production of vials and ampuls of easily oxidizable preparations intended for parenteral use.

Trace metals may originate from a number of sources, including the drug, solvent, container, or stopper and are a constant source of difficulty in preparing stable solutions of oxidizable drugs. Great care must be taken to eliminate these trace metals from labile preparations by thorough purification of the source of the contaminant or by chemically

complexing or binding the metal through the use of specialized agents that make it chemically unavailable for participation in the oxidative process. These agents are referred to as chelating agents and are exemplified by calcium disodium edetate and ethylenediamine tetra-acetic acid (EDTA).

Light can also act as a catalyst to oxidation reactions. As a photocatalyst, light waves transfer their energy (photons) to drug molecules, making the latter more reactive through increased energy capability. As a precaution against the acceleration of the oxidative process, sensitive preparations are packaged in light-resistant or opaque containers.

As most drug degradations proceed more rapidly with an increase in temperature, it is also advisable to maintain oxidizable drugs in a cool place. Another factor that could affect the stability of an oxidizable drug in solution is the pH of the preparation. Each drug must be maintained in solution at the pH most favorable to its stability. This, in fact, varies from preparation to preparation and must be determined on an individual basis for the drug in question.

In addition to oxidation and hydrolysis, other destructive processes such as polymerization, chemical decarboxylation, and deamination may occur in pharmaceutical preparations. However, these processes occur less frequently and are peculiar to only small groups of chemical substances.

19.6.16 Stability Testing

Drug and drug product stability testing during every stage of development is critical to the quality of the pharmaceutical product. Drug stability is important during preclinical testing and in clinical (human) trials to obtain a true and accurate assessment of the drug or drug product being evaluated. For a marketed drug product, assurance of drug stability is vital to the safety and effectiveness of the product when distributed and during the entire course of its shelf-life and use.

Product containers, closures, and other packaging features must be considered in stability testing. For instance, tablets or capsules packaged in glass or plastic bottles, blister packs, or strip packaging would require different stability test protocols. Drugs particularly subjected to hydrolysis or oxidative decomposition must be evaluated accordingly. In addition, parenteral and other sterile products must meet sterility test standards to ensure protection against microbial contamination. Any preservatives used must be tested for effectiveness in the finished product.

Drug products must meet stability standards for long-term storage at room temperatures and under conditions of relative humidity. Products are also subjected to accelerated stability studies as an indication of shelf-life stability. It is a Food and Drug Administration (FDA) requirement that if not submitted in the approved application, the first three postapproval production batches of a drug substance be placed on long-term stability studies and the first three postapproval production batches of a drug product be subjected to both long-term and accelerated stability studies.[7,8]

Drug instability in pharmaceutical formulations may be detected in some instances by a change in the physical appearance, color, odor, taste, or texture of the formulation. Whereas in other instances, chemical changes may occur which are not self-evident and may only be ascertained through chemical analysis. Scientific data pertaining to the stability of a formulation leads to the prediction of the expected shelf-life of the proposed product and, when necessary, to the redesignment of the drug (such as into more stable salt or ester form) and to the reformulation of the dosage form. Obviously, the rate or speed at which drug degradation occurs in a formulation is of prime importance. The study of the rate of chemical change and the way in which it is influenced by factors such

as, the concentration of the drug or reactant, the solvent employed, the conditions of temperature and pressure, and the presence of other chemical agents in the formulation is termed reaction kinetics.

In general, a kinetic study begins by measuring the concentration of the drug being examined at given time intervals under a specific set of conditions, including temperature, pH, ionic strength, light intensity, and drug concentration. The measurement of the drug's concentration at various time intervals reveals the stability or instability of the drug under the specified conditions with the passage of time. From this starting point, each of the original conditions may be altered on an individual basis to determine the influence that such changes make on the drug's stability. For example, the pH of the solution may be changed, whereas the temperature, light intensity, and original drug concentration remain as they were.

Data can be collected and may be presented graphically, by plotting the drug concentration as a function of time. From the experimental data, the reaction rate may be determined and a rate constant and half-life calculated.

Accelerated stability testing is the use of exaggerated conditions such as temperature, humidity, and light, to test the stability of drug formulations. Accelerated temperature stability studies, for example, may be conducted for 6 months at 40°C with 75% relative humidity. If a significant change occurs in the drug or drug product under these conditions, lower temperature and humidity may be used, such as 30°C and 60% relative humidity. The use of short-term accelerated studies is for the purpose of determining the most stable of the proposed formulations for a drug product. In stress testing, temperature elevations in 10° increments, higher than used in accelerated studies are employed until chemical or physical degradation occurs. Once the most stable formulation is ascertained, its long-term stability is predicted from the data generated from continuing stability studies. Depending on the types and severity of conditions employed, it is not unusual to maintain samples under exaggerated conditions of both temperature and varying humidity for periods of 6 to 12 months. Such studies lead to the prediction of shelf-life for a drug product.

19.6.17 Physical Stability and Signs of Degradation

Other signs of degradation of the specific dosage forms must be observed and reported. For the various dosage forms, this includes the following:[9]

- *Tablets*: Appearance (cracking, chipping, and mottling), friability, hardness, color, odor, moisture content, clumping, disintegration, and dissolution
- *Capsules*: Moisture tackiness, color, appearance, shape, brittleness, and dissolution
- *Oral solutions and suspensions*: Appearance, precipitation, pH, color, odor, redispersibility (suspensions), and clarity (solutions)
- *Oral powders*: Appearance, color, odor, and moisture
- *Metered-dose inhalation aerosols*: Delivered dose per actuation, number of metered doses, color, particle-size distribution, loss of propellant, pressure, valve corrosion, spray pattern, and absence of pathogenic microorganisms
- *Topical nonmetered aerosols*: Appearance, odor, pressure, weight loss, net weight dispensed, delivery rate, and spray pattern
- *Topical creams, ointments, lotions, solutions, and gels*: Appearance, color, homogeneity, odor, pH, resuspendibility (lotions), consistency, particle-size distribution, strength, and weight loss

- *Ophthalmic, nasal, and oral inhalation preparations*: Appearance, color, consistency, pH, clarity (solutions), particle size and resuspendibility (suspensions, ointments), strength, and sterility
- *Small-volume parenterals*: Appearance, color, particulate matter, dispersibility (suspensions), pH, sterility, pyrogenicity, and closure integrity
- *Large-volume parenterals*: Appearance, color, clarity, particulate matter, pH, volume and extractables (when plastic containers are used), sterility, pyrogenicity, and closure integrity
- *Suppositories*: Softening range, appearance, and dissolution or melting
- *Emulsions*: Appearance (as phase separation), color, odor, pH, and viscosity
- *Controlled-release membrane drug delivery systems*: Seal strength of the drug reservoir, decomposition products, membrane integrity, drug strength, and drug release rate

19.6.18 Expiration Dates and Beyond-Use Dates

Under usual circumstances, most manufactured products require a shelf-life of 2 or more years to ensure their stability at the time of patient consumption. Commercial products must bear an appropriate expiration date. This date identifies the time during which the product may be expected to maintain its potency and remain stable under the designated storage conditions. The expiration date limits the time during which the product may be dispensed by the pharmacist or used by the patient.

Prescriptions requiring extemporaneous compounding by the pharmacist do not require the extended shelf-life that commercially manufactured and distributed products do because they are intended to be used immediately on their receipt by the patient and used only during the immediate course of the prescribed treatment. However, these compounded prescriptions must remain stable and efficacious during the course of their use and the compounding pharmacist must employ formulative components and techniques, which results in a stable product.[10]

Whereas commercially manufactured products are required to possess an expiration date, compounded products are assigned a beyond-use date. There are numerous sources of information that can be used for determining an appropriate beyond-use date, such as chemical companies, manufacturers literature, laboratory data, journals, and published books on the subject. Generally, most pharmacists prepare or dispense small quantities of compounded products; recommend storage at room, cool, or cold temperatures; and use a conservative beyond-use date.

The guidelines published in the USP 27/NF 22 Section 795, Pharmacy Compounding, states that

> In the absence of stability information that is applicable to a specific drug and preparation, the following maximum beyond-use dates are recommended for nonsterile compounded drug precautions that are packaged in tight, light-resistant containers and stored at controlled room temperature unless otherwise indicated.[11]

- For nonaqueous liquids and solid formulations (for which the manufactured drug product is the source of active ingredient), the beyond-use date is not later than 25% of the time remaining until the product's expiration date or 6 months, whichever is earlier;
- When a USP or NF substance is the source of active ingredient, the beyond-use date is not later than 6 months;
- For water-containing formulations (prepared from ingredients in solid form), the beyond-use date is not later than 14 days when stored at cold temperatures;

- For all other formulations, the beyond-use date is not later than the intended duration of therapy or 30 days, whichever is earlier.[11]

These beyond-use date limits may be exceeded when there is supporting valid scientific stability information that is directly applicable to the specific preparation (i.e., the same drug concentration range, pH, excipients, vehicle, water content, etc.) [11]
When compounding on the basis of extrapolated or less concrete stability information, it is best for the pharmacist to keep the formulation simple and not take shortcuts. But, use the necessary pharmaceutical adjuvants to prepare the prescription.

19.7 Pharmaceutical Excipients

Although most drug substances in use today are unpalatable and unattractive in their natural state, modern pharmaceutical preparations present them to the patient as colorful, flavorful formulations attractive to the sight, smell, and taste. These qualities, which are the rule rather than the exception, have virtually eliminated the natural reluctance of many patients to take medications because of disagreeable odor or taste. In fact, the inherent attractiveness of today's pharmaceuticals has caused them to acquire the dubious distinction of being a source of accidental poisonings in the home, particularly among children who are lured by their organoleptic appeal.

Excipients are generally required to prepare a drug substance in final dosage form. For example, in the preparation of pharmaceutic solutions, one or more solvents are used to dissolve the drug substance. Flavors and sweeteners are used to make the product more palatable. Colorants are added to enhance product appeal. Preservatives may be added to prevent microbial growth. To prevent drug decomposition, chelating agents and stabilizers, and antioxidants may be used. In the preparation of tablets, diluents or fillers are commonly added to increase the bulk of the formulation. Binders are added to cause the adhesion of the powdered drug and pharmaceutic substances, antiadherents or lubricants to assist the smooth tableting process. Disintegrating agents are used to promote tablet break-up after administration, and coatings are used to improve stability, control disintegration, or to enhance appearance. Ointments, creams, and suppositories achieve their characteristic features due to the pharmaceutic bases that are utilized. Thus, for each dosage form, the pharmaceutic ingredients establish the primary features of the product, and contribute to the physical form, texture, stability, taste, and overall appearance.

There is some psychologic basis to drug therapy, and the odor, taste, and color of a pharmaceutical preparation can play a part. An appropriate drug will have its most beneficial effect when it is accepted and taken properly by the patient. The proper combination of flavor, fragrance, and color in a pharmaceutical product contributes to its acceptance.

19.7.1 Flavoring Pharmaceuticals

The flavoring of pharmaceuticals applies primarily to the liquid dosage forms intended for oral administration. The 10,000 taste buds, each having 60 to 100 receptor cells, interact with molecules dissolved in the saliva and produce a positive or negative taste sensation.[12] Obviously, medication in liquid form comes into immediate and direct contact with these taste buds. By the addition of flavoring agents to liquid medication, the disagreeable taste of drugs may be successfully masked. Drugs placed in capsules or prepared as coated tablets may be easily swallowed with avoidance of contact between the drug and the taste

buds. Tablets containing drugs that are not especially distasteful may remain uncoated and unflavored. Swallowing them with water usually is sufficient to avoid undesirable drug taste sensations. However, tablets of the chewable type such as certain antacid and vitamin products, which are intended for mastication in the mouth, usually are sweetened and flavored to receive better patient acceptance.

The formulation pharmacist is well acquainted with the taste characteristics of certain chemical types of drugs and strives to mask the unwanted taste through the appropriate use of flavoring agents. Although there are no dependable rules for unerringly predicting the taste sensation of a drug on the basis of its chemical constitution, experience permits the presentation of several observations. For instance, although we recognize and assume the salty taste of sodium chloride, the formulation pharmacist knows that all salts are not salty, but that their taste is a function of both the cation and anion. Salty tastes are evoked by sodium, potassium, and ammonium chlorides, and by sodium bromide. Potassium and ammonium bromides elicit simultaneous bitter and salty sensations. Predominantly bitter tastes result from potassium iodide and magnesium sulfate (epsom salt). In general, low-molecular-weight salts are salty, and higher-molecular-weight salts are bitter. With organic compounds, an increase in the number of hydroxyl groups (–OH) seems to increase the sweetness of the compound. In general, the organic esters, alcohols, and aldehydes are pleasant tasting, and since many of them are volatile, they also contribute to the odor and thus the flavor of preparations in which they are used. Many nitrogen-containing compounds are extremely bitter, especially the plant alkaloids (such as quinine), but certain other nitrogen-containing compounds are extremely sweet (such as aspartame). The medicinal chemist recognizes that even the simplest structural change in an organic compound can alter its taste. D-Glucose is sweet, but L-glucose has a slightly salty taste; saccharin is very sweet, but *N*-methyl-saccharin is tasteless.[13]

Thus, the predictability of the taste characteristics of a new drug is only speculative. However, it is soon learned, and the formulation pharmacist is then put to the task of increasing the drug's palatability in the environment of other formulative agents. The selection of an appropriate flavoring agent depends upon several factors, but primarily upon the taste of the drug substance itself. Certain flavoring materials are more effective than others in masking or disguising a particular bitter, salty, sour, or otherwise undesirable taste of medicinal agents. Although individuals' tastes and flavor preferences differ, cocoa-flavored vehicles are considered effective for masking the taste of bitter drugs. Fruit or citrus flavors are frequently used to combat sour or acid-tasting drugs, and cinnamon, orange, raspberry, and other flavors have been successfully used to make preparations of salty drugs more palatable.

Flavors can degrade as a result of exposure to light, temperature, headspace oxygen, water, enzymes, contaminants, and other product components so they must be carefully selected and checked for stability.

A general guide to using flavors is to start as follows. For water-soluble flavors, generally start at 0.2% for artificial and 1 to 2% for natural flavors. For oil soluble flavors, generally start at 0.1% in the finished product for artificial flavors and 0.2% for natural flavors. For solid or powdered flavors, one can generally start at 0.1% in the finished product for artificial flavors and 0.75% for natural flavors.

In flavoring liquid pharmaceutical products, the flavoring agent is added to the solvent or vehicle component of the formulation in which it is most soluble or miscible. That is, water-soluble flavorants are added to the aqueous component of a formulation and poorly water-soluble flavorants are added to the alcoholic or other nonaqueous solvent component of the formulation. In a hydroalcoholic or other multisolvent system, care must be exercised to maintain the flavorant in solution. This is accomplished by maintaining a sufficient level of solvent in which the flavorant is soluble.

19.7.2 Sweetening Pharmaceuticals

In addition to sucrose, a number of other sweetening agents have been utilized in foods and pharmaceuticals over the years, including dextrose, mannitol, sorbitol, aspartame, saccharin, and others. Some sweeteners, such as sucrose, aid in preserving the product.

19.7.3 Coloring Pharmaceuticals

Coloring agents are used in pharmaceutical preparations for purposes of esthetics. A distinction should be made between agents that have inherent color and those that are employed as colorants. Certain agents such as sulfur (yellow), riboflavin (yellow), cupric sulfate (blue), ferrous sulfate (bluish green), and cyanocobalamin (red), have inherent color and are not thought of as pharmaceutical colorants in the usual sense of the term.

Although most pharmaceutical colorants in use today are of synthetic origin, a few are obtained from natural mineral and plant sources. For example, red ferric oxide is mixed in small proportions with zinc oxide powder to prepare calamine, giving the latter its characteristic pink color, which is intended to match the skin tone upon application.

The synthetic coloring agents used in pharmaceutical products were first prepared in the middle of the 19th century from principles of coal tar. Coal tar (pix carbonis), a thick, black, viscid liquid, is a by-product in the destructive distillation of coal. Its composition is extremely complex, and many of its constituents may be separated by fractional distillation. Among the products obtained are anthracene, benzene, naphtha, creosote, phenol, and pitch. About 90% of the total dyes used in the products that FDA regulates are synthesized from a single, colorless derivative of benzene, called aniline. These aniline dyes are also known as synthetic organic dyes or as "coal tar" dyes, since aniline was originally obtained from bituminous coal. Aniline dyes today come mainly from petroleum.

A colorant becomes an integral part of a pharmaceutical formulation, and its exact quantitative amount must be reproducible each time the formulation is prepared, or else the preparation would have a different appearance from batch to batch. This requires a high degree of pharmaceutical skill, for the amount of colorant generally added to liquid preparations, the range can be between 0.0005 and 0.001% depending upon the colorant and the depth of color desired. Because of their color potency, dyes are generally added to pharmaceutical preparations in the form of diluted solutions rather than as concentrated dry powders. This permits greater accuracy in measurement and more consistent in color production.

In addition to liquid dyes in the coloring of pharmaceuticals, lake pigments may also be used. Whereas a chemical material exhibits coloring power or tinctorial strength when dissolved, pigment is an insoluble material, which colors by dispersion. An FD&C lake is a pigment consisting of a substratum of alumina hydrate on which the dye is absorbed or precipitated. Having aluminum hydroxide as the substrate, the lakes are insoluble in nearly all solvents. FD&C lakes are subject to certification and must be made from dyes that have been previously certified. Lakes do not have a specified dye content and range from 10 to 40% pure dye. By their very nature, lakes are suitable for coloring products in which the moisture levels are low.

In the preparation of capsules, various colored empty gelatin capsule shells may be used to hold the powdered drug mixture. Many commercial capsules are prepared with capsule bodies of one color and a different colored capsule cap, resulting in a two-colored capsule. This makes certain commercial products even more readily identifiable than solid colored capsules. For powdered drugs, dispensed as such or compressed into tablets, a generally larger proportion of dye is required (about 0.1%) to achieve the desired hue than with liquid preparations.

Both dyes and lakes have application in the coloring of sugar-coated, film-coated, and direct-compression tablets, and pharmaceutical suspensions and other dosage forms.[14] Traditionally, sugar-coated tablets have been colored with syrup solutions containing varying amounts of the water-soluble dyes, starting with very dilute solutions, working up to concentrated color syrup solutions. As many as 30 to 60 coats are not uncommon. Using the FD&C lakes, fewer color coats are used. Appealing tablets have been made with as few as 8 to 12 coats using lakes dispersed in syrup. Water-soluble dyes in aqueous vehicles or lakes dispersed in organic solvents may be effectively sprayed on tablets to achieve attractive film coatings. There is continued interest today in chewable tablets, due to the availability of many direct-compression materials such as dextrose, sucrose, mannitol, sorbitol, and spray-dried lactose. The direct-compression colored chewable tablets may be prepared utilizing 1 pound of lake per 1000 pounds of tablet mix. For aqueous suspensions, FD&C water-soluble colors or lakes may be satisfactory. In nonaqueous suspensions, FD&C lakes are necessary. The lakes are added either to the aqueous or nonaqueous phase, generally at a level of 1 pound of color per 1000 pounds of suspension, and require homogenizing or mechanical blending to achieve uniform coloring.

In most cases, ointments, suppositories, ophthalmic, and parenteral products assume the color of their ingredients and do not contain color additives. In addition to esthetics and the certification status of a dye, a formulation pharmacist must select the dyes to be used in a particular formula on the basis of the physical and chemical properties of the dyes available. Of prime importance is the solubility of a prospective dye in the vehicle to be used for a liquid formulation or in a solvent to be employed during a pharmaceutical process (such as when the dye is sprayed on a batch of tablets). In general, most dyes are broadly grouped into those that are water-soluble and those that are oil-soluble; few, if any, dyes are both.

Another important consideration when selecting a dye for use in a liquid pharmaceutical is the pH and pH stability of the preparation to be colored. Dyes can change color with a change in pH, and a dye must be selected for a product so that any anticipated pH change will not alter the color during the usual shelf-life. The dye must also be chemically stable in the presence of other formulative ingredients and must not interfere with the stability of other agents. To maintain their original colors, FD&C dyes must be protected from oxidizing agents, reducing agents (especially metals as iron, aluminum, zinc, and tin), strong acids and alkalis, and excessive heating. Dyes must also be reasonably photostable; that is, they must not change color when exposed to light of anticipated intensities and wavelengths under the usual conditions of shelf storage. Certain medicinal agents, particularly those prepared in liquid form, must be protected from light to maintain their chemical stability and their therapeutic effectiveness. These preparations are generally maintained and dispensed in dark amber or opaque containers. For solid dosage forms of photolabile drugs, a colored or opaque capsule shell may actually enhance the drug's stability by shielding out light rays.

19.7.4 Sterilization and Preservation

In addition to the stabilization of pharmaceutical preparations against chemical and physical degradation due to changed environmental conditions within a formulation, certain liquid and semisolid preparations also must be preserved against microbial contamination.

Although some types of pharmaceutical products, like ophthalmic and injectable preparations, are sterilized by physical methods, including autoclaving, dry heat, or by bacterial filtration during their manufacture, many of them additionally require the presence of an antimicrobial preservative to maintain their aseptic condition throughout the period of their storage and use. Other types of preparations that are not sterilized during their

preparation but are particularly susceptible to microbial growth because of the nature of their ingredients are protected by the addition of an antimicrobial preservative. Preparations that provide excellent growth media for microbes are most aqueous preparations, especially syrups, emulsions, suspensions, and some semisolid preparations, particularly creams. Certain hydroalcoholic and most alcoholic preparations may not require the addition of a chemical preservative when the alcoholic content is sufficient to prevent microbial growth. Generally, 15 to 18% alcohol will prevent microbial growth in an acid media and 18% in an alkaline media. Most alcohol-containing pharmaceuticals such as elixirs, spirits, and tinctures are self-sterilizing and do not require additional preservation. The same would apply to other pharmaceuticals on an individual basis, which by virtue of their vehicles or other formulative agents, may not permit the growth of microorganisms.

19.7.4.1 Preservative Selection

When experience or shelf-storage experiments indicate that a preservative is required in a pharmaceutical preparation, its selection is based on many cross considerations including some of the following:

- The preservative must prevent the growth of the type of microorganisms considered the most likely contaminants of the preparation being formulated.
- The preservative must be sufficiently soluble in water to achieve adequate concentrations in the aqueous phase of a two or more phase system. The proportion of preservative remaining undissociated at the pH of the preparation must be capable of penetrating the microorganisms and destroying its integrity.
- The required concentration of the preservative must not affect the safety or comfort of the patient when the pharmaceutical preparation is administered by the usual or intended route (such as nonirritating, nonsensitizing, and nontoxic).
- The preservative must have adequate stability and should not be reduced in concentration due to chemical decomposition or volatilization during the desired shelf-life of the preparation.
- The preservative must be completely compatible with all other formulative ingredients and does not interfere with them, nor do they interfere with the effectiveness of the preservative agent.
- The preservative must not adversely affect the preparation's container or the closure.

19.7.5 General Preservative Considerations

Microorganisms involved include molds, yeasts, and bacteria, with the latter generally favoring a slightly alkaline medium and the others an acid medium. Although, few microorganisms can grow below a pH of 3 or above a pH of 9, most aqueous pharmaceutical preparations are within the favorable pH range and therefore must be protected against microbial growth. To be effective, a preservative agent must be dissolved in sufficient concentration in the aqueous phase of a preparation. Further, only the undissociated fraction or molecular form of a preservative possesses preservative capability, because the ionized portion is incapable of penetrating the microorganism. Thus the preservative selected must be largely undissociated at the pH of the formulation being prepared. Acidic preservatives like benzoic, boric, and sorbic acids, are more undissociated and thus more effective as the medium is made more acidic. Conversely, alkaline preservatives are less effective in acid or neutral media and more effective in alkaline media. Thus, it is meaningless to suggest preservative effectiveness at specific concentrations unless the pH of the

system is mentioned and the undissociated concentration of the agent is calculated or otherwise determined. Also, if formulative materials interfere with the solubility or availability of the preservative agent, its chemical concentration may be misleading, because it may not be a true measure of the effective concentration. Many incompatible combinations of preservative agents and other pharmaceutical adjuncts have been identified and more will probably be found in the future as new preservatives, pharmaceutical adjuncts, and therapeutic agents are combined in new ways. It is important for the formulation pharmacist to examine all formulative ingredients since one affects the other. This assures the formulative pharmacist to assure himself that each agent is free to do the job for which it was intended in the formulation. In addition, the preservative must not interact with a container such as a metal ointment tube, plastic medication bottle, or with an enclosure such as a rubber or plastic cap or liner. Such an interaction could result in the decomposition of the preservative or the container closure, or both, with resultant product decomposition and contamination. Appropriate tests should be devised and conducted to insure against this type of preservative interaction.

19.8 Regulations and Guidelines Affecting Pharmaceutical Compounding

Documents of special importance in providing guidelines and standards for pharmaceutical compounding include the National Association of Boards of Pharmacy's "Good Compounding Practices Applicable to State Licensed Pharmacies;" the USP 27/NF 22 Chapter 795, "Pharmacy Compounding — Nonsterile Preparations;" and Chapter 797, "Pharmacy Compounding — Sterile Preparations;" as well as numerous other portions of the USP/NF.

19.8.1 Good Compounding Practices Applicable to State-Licensed Pharmacies

The following Good Compounding Practices (GCPs) are meant to apply only to the compounding of drugs by state-licensed pharmacies. These GCPs are divided into eight subparts.

19.8.1.1 Subpart A: General Provisions

General provisions contain recommendations that are considered the minimum current good compounding practices for the preparation of drug products by state-licensed pharmacies for dispensing or administration to humans or animals. The document contains definitions of compounding and manufacturing.

19.8.1.2 Subpart B: Organization and Personnel

Organization and personnel describe that in the dispensing of all prescriptions, the pharmacist has the responsibility and authority to inspect and approve or reject all components, drug product containers, closures, in-process materials, and labeling. It also has the authority to prepare and review all compounding records to assure that no errors have occurred in the compounding process. The pharmacist is also responsible for the proper maintenance, cleanliness, and use of all equipment used in prescription compounding practice.

19.8.1.3 Subpart C: Drug Compounding Facilities

Drug compounding facilities discuss that pharmacies engaging in compounding shall have a specifically designated and adequate area (space) for the orderly placement of equipment and materials to be used to compound medications. The drug compounding area for sterile products shall be separate and distinct from the area used for the compounding or dispensing of nonsterile drug products. The area(s) used for the compounding of drugs shall be maintained in a good state of repair. If sterile (aseptic) products are being compounded, conditions set forth in the NABP "Model Rules for Sterile Pharmaceuticals" must be followed. If radiopharmaceuticals are being compounded, conditions set forth in the NABP "Model Rules for Nuclear/Radiologic Pharmacy" must be followed.

If drug products with special precautions for contamination, such as penicillin, are involved in a compounding operation, appropriate measures, including either the dedication of equipment for such operations or the meticulous cleaning of contaminated equipment prior to its return to inventory, must be used to prevent cross-contamination.

19.8.1.4 Subpart D: Equipment

Equipment explains that equipment used in the compounding of drug products shall be of appropriate design, adequate size, and suitably located to facilitate operations for its intended use and for its cleaning and maintenance. Equipment used in the compounding of drug products shall be of suitable composition so that surfaces that contact components, in-process materials, or drug products shall not be reactive, additive, or absorptive so as to alter the safety, identity, strength, quality, or purity of the drug product beyond that desired.

19.8.1.5 Subpart E: Control of Components and Drug Product Containers and Closures

Control of components and drug product containers and closures requires that components, drug product containers, and closures, used in the compounding of drugs shall be handled and stored in a manner to prevent contamination. Bagged or boxed components of drug product containers and closures used in the compounding of drugs shall be stored off the floor in such a manner so as to permit cleaning and inspection.

19.8.1.6 Subpart F: Drug Compounding Controls

Drug compounding controls state that there shall be written procedures for the compounding of drug products to assure that the finished products have the identity, strength, quality, and purity they purport or are represented to possess. Such procedures shall include listing of the components (ingredients), their amounts (in weight or volume), the order of component addition, and a description of the compounding process. All equipment and utensils and the container or closure system, relevant to the sterility and stability of the intended use of the drug, shall be listed. These written procedures shall be followed in the execution of the drug compounding procedure.

19.8.1.7 Subpart G: Labeling Control of Excess Products

Labeling control of excess products explains that in the case where a quantity of a compounded drug product in excess of that to be initially dispensed in accordance with Subpart A is prepared, the excess product shall be labeled or documentation referenced

with the complete list of ingredients (components), the preparation date, and the assigned expiration date based upon professional judgment, appropriate testing, or published data. It shall also be stored and accounted for under conditions dictated by its composition and stability characteristics (such as in a clean, dry place on a shelf or in the refrigerator) to ensure its strength, quality, and purity.

19.8.1.8 Subpart H: Records and Reports

Records and reports require that any procedures or other records required to be maintained in compliance with these Good Compounding Practices shall be retained for the same period of time as each State requires for the retention of prescription files.

19.8.2 U.S. Pharmacopeia — National Formulary

The following are summaries of the lengthy Chapter 795, "Pharmacy Compounding — Nonsterile Preparations," and Chapter 797, "Pharmacy Compounding — Sterile Preparations, in the USP/NF."

USP Chapter 795 "Pharmacy Compounding — Nonsterile Preparations" is divided into the following nine sections:

- Compounding environment
- Stability of compounded preparations
- Ingredient selection and calculations
- Checklist for acceptable strength, quality, and purity
- Compounded dosage forms
- Compounding process
- Compounding records and documents
- Quality control
- Patient counseling

The compounding environment section describes guidelines for the facilities and equipment that should be available, calibrated, maintained and, used in a compounding pharmacy. The stability section has been cited previously in this chapter, in part, to provide guidelines for beyond-use dates to be placed on compounded preparations. The selection of ingredients has also been previously discussed in this chapter. The checklist for the USP/NF hallmarks of standards of acceptable strength, quality, and purity is presented in a series of questions to be answered. Examples of compounded dosage forms are discussed along with some precautionary statements as appropriate. A step-by-step presentation on the compounding process is outlined to ensure uniformity of activities during each preparation. Documentation is described for the *Formulation Record*, the *Compounding Record*, and the *Material Safety Data Sheets* (MSDS) files that should be maintained. The section ends with various aspects for patient counseling involving the proper use, storage, and evidence of instability of the compounded preparation.

USP Chapter 797 Pharmacy Compounding — Sterile Preparations is divided into the following ten categories:

- Responsibility of the dispensing pharmacist
- Risk levels
- Validation
- Low-risk operations

- High-risk operations
- Environmental quality and control, finished product release checks, and tests
- Storage and expiration dating
- Maintaining product quality and control after it leaves the pharmacy
- Patient or caregiver training
- Patient monitoring and complaint system

The compounding pharmacist dispensing any sterile drug preparation is responsible for ensuring that the product has been prepared, labeled, controlled, stored, dispensed, and distributed properly. Low- and high-risk levels of sterile products compounding are defined with examples of each. Validation of the sterilization and aseptic processing procedures are described as related to both personnel, facilities or equipment, and processes. Low- and high-risk operations are described along with the validations required for each. Environmental quality and control procedures for the work area and personnel along with suggested Standard Operating Procedures (SOPs) and an example of an environmental monitoring program are described.

Tests and procedures for the finished product are described with the guidelines that only those products that are free from defects and meeting all quality specifications will be distributed. Guidelines are discussed for preparation, storage, and beyond-use dating after the preparation leaves the pharmacy. After the preparation leaves the pharmacy, the caregiver or patient should receive training to ensure understanding and compliance with the storage, handling, and administration of the preparation. The various aspects of the training program are outlined in this chapter. Also included is the recommendation for written policies and procedures for the monitoring of patients using home-use sterile drug products and the handling and reporting of adverse events. It is evident in this chapter that the responsibility of the compounding pharmacist ranges from the activities involved in the compounding of the product through its proper storage, distribution, use, and disposal.

19.9 Summary

Pharmaceutics, pharmaceutical manufacturing, and pharmaceutical compounding share a symbiotic relationship to provide quality pharmaceuticals to patients. Manufacturers can most efficiently provide quality products on a large scale to the masses and pharmaceutical compounding can provide specific preparations for individual patients.

The use of pharmaceutics and most of the pharmaceutical or physicochemical factors involved in formulation of these products are identical. One exception, however, is the shorter beyond-use dates assigned to compounded preparations as compared to expiration dates for manufactured products. A second notable exception is that commercially manufactured products are developed over several years whereas an extemporaneously compounded preparation may be required for a specific patient in a matter of minutes or hours.

Pharmaceutical manufacturing will continue to grow as more new, innovative, and novel dosage forms and drug delivery systems become routine. Pharmaceutical compounding will continue to grow and become an even more important part of pharmacy practice in the future as individualization of patient care becomes more important. Working together, the needs of most patients will be satisfactorily addressed.

References

1. U.S. Pharmacopeia 25, Rockville, MD, U.S. Pharmacopeial Convention, 2001.
2. National Association of Boards of Pharmacy, Good compounding practices applicable to state-licensed pharmacies, in *Model State Pharmacy Act and Model Rules of the National Association of Boards of Pharmacy*, Park Ridge, IL, C1, 1993.
3. The Food and Drug Administration Modernization Act of 1997, Pub. Law 105–115, 127, 1997.
4. National Association of Boards of Pharmacy, *Good Compounding Practices Applicable to State Licensed Pharmacies*, Park Ridge, IL, 1993.
5. Allen, L.V., Jr., A history of pharmaceutical compounding, *Secundum Artem*, 11, 3, 2003.
6. Brange, J., Langkjaer, L., Havelund, S., and Vϕlund, A., Chemical stability of insulin. Hydrolytic degradation during storage of pharmaceutical preparations, *Pharm. Res.*, 9, 715, 1991.
7. Sheinin, E.B., ICH guidelines, History, Present Status, Intent, presented at the International Good Manufacturing Practices Conference, Athens, GA, 1998.
8. Rothman, B., Stability is the Issue, presented at the International Good Manufacturing Practices Conference, Athens, GA, 1998.
9. Poole, J.W., Preformulation, FMC Corporation, 1982.
10. American Society of Hospital Pharmacists, ASHP technical assistance bulletin on compounding nonsterile products in pharmacies, *Am. J. Hosp. Pharm.*, 51, 1441–1448, 1994.
11. U.S. Pharmacopeia 27, Rockville, MD, U.S. Pharmacopeial Convention, 2001.
12. Lewis, R., When smell and taste go awry, *FDA Cons.*, 25, 29, 1991.
13. Hornstein, I. and Teranishi, R., The chemistry of flavor, *Chem. Eng. News*, 45, 92,1967.
14. Anonymous, Colorants for drug tablets and capsules, *Drug Cosmet. Ind.*, 133, 44, 1983.

20

Late Stage and Process Development Activities

Charles G. Smith

In this chapter, and several that follow, we have chosen to bring together those activities that really are a continuum from the programs mentioned above but those, once again, require special expertise and experience in drug development to be completed in timely fashion for new drug application (NDA) and overseas regulatory filings. Under "late stage," I include the following processes as being necessary to be brought to fruition on a time-sensitive, or critical-path, basis:

1. The completion of Phase III clinical trials in the United States (large scale for use with the U.S. Food and Drug Administration (FDA) and conducted globally in many companies) and the writing of the clinical document for the NDA, complete with the biostatistical section.
2. In conjunction with international colleagues, the decision as to when sufficient clinical data from the worldwide development program have been amassed to permit filing for product license applications overseas should be made. (If the studies are designed correctly, this date may precede the U.S. filing date by months. This approach can make filing overseas significantly more efficient than it is in the Untied States and every effort should be made to take advantage of this fact by proper coordination with the international group from inception of the program.) Great care must be taken to insure that the program for the U.S.–NDA is not compromised in any way by overseas filings. If any significant risk to the NDA is perceived, early filings outside the United States are ill advised.
3. Pharmacology, biochemistry, and drug metabolism testing should all be completed.
4. All phases of the toxicology studies, including carcinogenicity studies, where required, be completed and final reports issued.
5. Pharmaceutics work should be completed, including stability studies on the final formulation that can be utilized worldwide.
6. Large-scale processes for preparing bulk drug must be in place, with all the procedures required by the FDA and other sophisticated regulatory bodies satisfied. In the United States, the FDA must inspect the facility in which the drug will ultimately be manufactured (if it has not already been inspected) to give final NDA approval.
7. Complete worldwide patent filings must be reviewed and appropriate action taken to insure adequate protection for worldwide commercialization.
8. Plans to publish a series of scientific and clinical papers should be initiated.

In the early stages of the drug-development process, the amounts of bulk drug substance required for pharmacology studies, early pharmaceutics investigations, and early toxicology are usually prepared in the chemistry department (or the chemistry section of the microbiology department if the product is derived from a fermentation) or by the biochemistry or biotechnology laboratory groups using small-scale equipment. Relatively early in the process, when a drug candidate is undergoing pharmacology and early toxicology studies, kilogram quantities (of the usual synthetic compounds) may be needed to continue the toxicology and pharmaceutics development programs. In addition to preparing the kilogram quantities of bulk drug required for these investigations, the so-called process development group must undertake studies to derive a commercially viable process to prepare bulk material in hundreds or thousands of kilogram quantities (again, synthetic compounds, not biotechnology products) for ultimate marketing. Process development groups can specialize in straight chemical synthetic operations, biotechnology production processes, or isolation procedures (to extract the drug candidate(s) from fermentation beers, plants, or blood plasma). In the usual pharmaceutical operation, small laboratory (bench) scale will be used to prepare between 10 and 100 g of material (this excludes biotechnology processes) and intermediate pilot plant scale will be used to prepare amounts ranging from hundreds of grams to amounts from 1 to 10 kg. Full-scale pilot plant or production facilities will generally be used to prepare quantities greater than 100 kg.

In the case of small companies, all production steps will be carried out by a contract organization.

Biotechnology products require different kinds of equipment for both small- and large-scale operations and this subject will be covered in detail in later chapters.

The problem of "scale-up" (e.g., moving from small-scale synthesis at the laboratory bench to several hundred- or thousand-liter capacity tanks) is a very real one. In my experience, the best way to insure prompt scale-up, with minimal finger pointing and arguing over whose "fault" it is if the process does not work at the larger scale, is to involve the process development group early in the drug development program. Indeed, scientists from process development should be represented on project teams so that they are fully *au courant* with drug candidates moving through the system. Members of the chemistry department, while making 50- to 100-g quantities of material at the bench, should invite their colleagues from process development to observe the process and to give it some thought with respect to ultimate scale-up. The first run in the pilot plant should be carried out with appropriate research laboratory scientists collaborating with the development group from beginning to end.

If pharmacology and toxicology investigations continue to give a new molecule the green light to move toward the clinic, process development should become involved in handling the compound and in making the supplies that will be required for more extensive toxicology, pharmaceutics, and clinical investigations. Many questions must be considered by process development that are not problems when small-scale syntheses are employed. For example, not only must the arithmetic or logarithmic relationships between volumes in a flask and volumes in a reactor or on a large column be evaluated, but also the relative rates of dissipation of heat from the reaction vessels, adequate mixing of gases and solvents, physical ability to pack a column uniformly after a certain diameter is reached, etc. must be addressed. In addition, there are real questions of cost and safety in the use of certain solvents that may be totally acceptable at the laboratory scale, but are either too expensive or too dangerous (e.g., ether) at the pilot plant or production scale of operation. Through the pilot plant experience of producing materials for toxicology and pharmaceutics investigations, the process development team gets a very good "feel" about the molecules and the problems that will be associated with their ultimate scale-up to full plant operation. By the time the NDA is ready to be filed, the company should have a commercially feasible

process in hand that will allow the manufacture of those quantities of material that are needed for ultimate marketing.

With biotechnology products, the entire process must be carefully overseen/monitored by the scientists who perform the genetic engineering operations. Interactions with third-party manufacturing operations must be frequent.

If the product is to be marketed in the United States, the plant must meet the FDA's good manufacturing practices (GMP) requirements (which now apply in most of the developed countries). If other products are being manufactured in a given plant for sale in the United States, it is not a certainty that the FDA will inspect the plant for the production of each new compound that is to be produced therein. It is virtually certain that the plant will be inspected, however, if the company has not been previously cleared by the FDA for manufacturing or sale in the United States. Inspection is also likely if the new process represents a significant deviation from the processes that have been carried out in the plant in the past. Requests to the FDA for plant inspection should be made as early in the NDA cycle as the law permits as scheduling the actual date for inspection can be a problem.

In those cases where long lead times for production, or a long processing time or cycle, result in large inventories (e.g., in the plasma fractionation business or in the extraction of certain plants that are seasonal and can only be harvested at certain times of the year), lead times for manufacturing the batches that will be used for clinical studies and market introduction become extremely critical. It is imperative that an adequate project review system be in place that establishes exactly where the bulk material for sale will be made and the availability of all raw materials, in FDA approved form, so that there is no default on the part of the pilot plant or manufacturing group in supplying bulk material at the time the drug is finally approved for sale. For ultimate FDA approval, the production group has to manufacture several batches of finished product and conduct shelf-life stability studies thereon to show that the product can be reproducibly manufactured in production facilities.

While the process development group scales up the bulk drug manufacturing process, the pharmaceutics development department performs a similar function with respect to the final dosage form. Research and development pharmacy groups must maintain contact with their colleagues in the production division to insure smooth flow of information so that large-scale production of acceptable final formulation will be insured after NDA approval. The period between submission of an NDA and ultimate marketing of a product is one that can be rather touchy in the area of bulk and finished formulation manufacturing. The reason for this statement is the fact that production does not want to make large amounts of material, to which an expiration date must be affixed months before the drug is actually approved for marketing. On the other hand, once approval is received, the marketing people do not want to hear about delays in producing their new product, for which they have been waiting for so long. During the "incubation" period at the FDA, which ranges from 15 to 18 months for a very few NDAs to 2.5 to 3 years or longer for many NDAs, manufacturing will stand ready to produce the material and stay in close touch with the drug regulatory group awaiting their advice on the timing of FDA approval. Once an "approvable letter" is received from the FDA, the company is assured that the drug will be approved when the company and the FDA agree on the final labeling. By that time, the production division should be under way preparing the first batches for marketing.

Sterile products for injection represent a particular challenge for the pharmaceutics development group. To prepare injectables, the pharmacists need not only sterile rooms in which to work at the laboratory, pilot plant, and production scales of operation, but they also require pyrogen-free water. Pyrogens are impurities, generally originating with

microbial contaminants, which cause spiking fevers in humans and certain other animal species. The water to be used in the preparation of sterile products must be tested (usually by the *in vitro* Limulus test or the standard pyrogen test in rabbits) at specified intervals to insure that it remains pyrogen-free. The tedious requirements for air sterility, pyrogen-free water, etc. that go along with GMP sterile manufacturing must be maintained throughout this operation.

In companies that are large enough to have their own clinical supplies preparations group, an appropriate facility with separate air supply, separation of rooms for handling different drug substances, and other regulations pertaining to the preparation of clinical supplies must be provided. Most medium-size or small companies will undoubtedly find it advantageous to contract clinical supplies manufacturing with a third party that provides such a service.

The "late stage" operations in the field of drug discovery and development are, as one would expect, primarily focused on the "development" side of the coin. The very important interface at the point where operations such as bulk material preparation, high-volume clinical supplies, extensive editing and reviewing of documents to be filed with the regulatory agencies, should bring the "research" group into close operating contact with the "development" people rather early in the game. The creation of "area teams," composed of people who must communicate regularly and frequently on matters of due dates and availability of supplies, is very important. Major delays can develop in a program when people at interfaces are not in tune with the time/scope aspect of the development job. The better informed all workers in the project are, the more likely that important dates will be met and the research teams and management will experience fewer disappointments in the rate of progress of a new drug candidate. One of my superiors once said to me, "We need a new drug fast and we need it right but, most of all, it *must* be right!" In the push for new products and the impossibility of always making accurate predictions, research and development must work closely, making their best "educated guesses" about achieving goals and communicating bad, as well as good, news to the management on a timely basis. Such operations will better insure that appropriate adjustments in plans and schedules can be made in all participating groups in an efficient manner.

Section IV

Clinical Development

21

Contract Research Organizations: Role and Function in New Drug Development

F. Richard Nichol

CONTENTS

21.1 Introduction

Most of the new pharmaceuticals, biologicals, medical devices, and selected diagnostics require clinical trials to generate data supporting their safety and efficacy in humans. Safety and efficacy data are mandatory for regulatory approval leading to commercialization in advanced countries. Clinical trials are the backbone of the human therapeutic development process and, as such, are regulated by government agencies to ensure clinical effectiveness and patient safety. The expansion of the Pure Food and Drug Act in

the United States in 1962 to contain regulations requiring rigorous proof of safety and efficacy was the genesis of modern product development methodologies in the United States. As these methods developed, they were the principal influences for the evolution of international standards and regulations necessary for pharmaceutical and medical device product approval in advanced countries.

From 1962 until 1975, commercial product developers, called sponsors, performed the majority of regulation-driven product development functions internally. Occasionally, agreements were struck with academicians and research institutes. For example, Stanford Research Institute (Menlo Park, CA, U.S.A.), Battelle Memorial Laboratories (Columbus, OH, U.S.A.), and several commercial animal toxicology firms performed specific activities necessary for product development, testing and registration. These included difficult chemical syntheses or formulations, clinical pharmacology, and toxicology. This involved medical specialists dosing normal human subjects and consultation with academic statisticians on clinical study design.

In 1975, the first commercial entity was formed to provide multicenter, outpatient clinical trial contracting services in the United States.[1] This was the beginning of the modern Contract Research Organization (CRO) industry. It is now composed of thousands of domestic and international firms involved in medical product evaluations in humans.[2] The total annual expense globally for clinical trials and related functions necessary for commercial, government, and academic therapeutic product development is estimated at US $50 billion, of which about US $15 billion is outsourced to the contract services market.

The current CRO market consists of four general categories of commercial contractors. The first group is composed of full-service, multinational, and publicly traded firms with revenues and market valuations in billions of dollars. The ten largest CROs account for approximately US $9 billion in combined annual revenues. The companies in the other three groups include midsize CROs, niche or specialty contractors, and Site Management Organizations (SMOs). These groups contribute another $6 billion annually. Since several large CROs offer preclinical animal pharmacology and toxicology services, and some also have trial-centric reference labs, those revenues are included in the $15 billion market estimate.

Since the mid-1990s, many new niche contractors were established by capitalizing on advances in information technology (IT). The therapeutic product development process includes clinical trials that are extremely complex, and many functions can be dramatically improved by applying customized IT applications.

Opportunities for improving the product registration process, especially those pertaining to clinical trials, have attracted large IT companies, such as Oracle (Redwood Shores, CA, U.S.A.), IBM (White Plains, NY, U.S.A.), and Siebel Systems (San Mateo, CA, U.S.A.). Additionally, Accenture (New York, NY, U.S.A.), Capgemini (Paris), and many midsize consultancies, such as First Consulting Group (Long Beach, CA, U.S.A.), have established clinical trial offerings involving systems integration and related products and services.

Virtually all clinical trials designed to support new therapeutic product registration require comprehensive laboratory tests performed on specimens from test subjects and patients. The reference laboratory business has a major presence in this market and companies such as Quest Diagnostics (Teterboro, NJ) and Laboratory Corporation of America (Burlington, NC, U.S.A.) are expected to grow impressively in revenue and profit, partly due to the incorporation of new biomarkers in clinical trial protocols. These markers, when validated, are part of the pharmacogenomic science that is emerging as an important component of a progressive regulatory strategy with the Food and Drug Administration (FDA) and other oversight agencies.

21.2 Role of the Contract Research Organization (CRO) Industry

The principal role of the CRO is to contribute to the enhancement of the financial performance of its customers (Table 21.1). Improvement in the efficiency of their clients' product development activities provides benefits to regulators, physicians and their patients. These include reduced cycle time, availability of external capacity, increased quality of trial management, and higher quality clinical data, all of which can result in lower labor costs. In late stage peri-approval trials, CROs may also play a role in developing safety profiles in large patient populations that are exposed to a new product after it is approved for manufacture and sale. Such trials can also assist the revenue growth of a new product.

The development required for approval of a new medication is lengthy (4 to 8 years), expensive ($200 to 800 million), risky, and subject to significant government regulations. This segment involves progressive clinical research, usually involving up to 3000 or 4000 research patients, who must be identified, screened for compliance with the strict criteria contained in the clinical research protocol, properly enrolled in the study, and carefully managed throughout the course of the trial. Clinical trials represent the bulk of product development expense. Contract Research Organizations provide services designed to improve and accelerate these trials, resulting in superior return on investment for the new product.

In the current, price-sensitive environment, which is the hallmark of modern healthcare and certainly new drug development, expenses necessary to support clinical research are a major concern for sponsors. Both fixed and variable costs can be managed effectively compared to complete in-house development budgets. As product development costs are a significant element influencing pharmaceutical pricing, the role of innovative CROs is expected to grow.

To help shorten registration time and reduce costs, CROs can provide benefits to sponsors listed below. Since each of these factors can have a major beneficial impact on the financial performance of the sponsor, it may be helpful to describe the impact of each on the development process. If a sponsor elects to combine several of these elements in an effective fashion, the resulting positive leverage on the trial process can be substantial.

21.2.1 Increased External Capacity

Considering the risk of modern clinical research and development of new therapeutic entities, sponsors frequently look to CROs to provide increased capacity for projects that need significant additional manpower. However, capacity issues are also a major concern for CROs, as they must juggle resources for new contracts, cancelled programs, frequently unpredictable rates of patient enrollment, and unexpected regulatory holds. The interaction

TABLE 21.1

Major Benefits to Sponsors of Using Contract Research Organizations

Increased external capacity
Reduced cycle time
Accelerating the learning curve
Economies of scale
Rapid go/no-go decisions
Innovation

between the sponsor and CRO for potential future projects can be delicate at best. Reserving capacity for future studies is very expensive for sponsors, and that approach has not proven to be popular in this contracting environment. To help insure rapid initiation of trials, the execution of master contract agreements can be helpful to moving forward swiftly, since legal issues, regulatory requirements, pricing, and change-order procedures can be codified prior to the initiation of the trial.

Since many promising compounds fail early in the product development cycle, leveraging skilled external capacity provides therapeutic product developers with the flexibility to ramp up or down this capacity to achieve economic efficiency. Compounds in developmental pipelines can be increased dramatically by in-licensing of new product candidates, a strategy employed by most large companies to support essential financial growth. CROs, SMOs, and niche service providers play important roles in the constant balancing act of resource allocation in the modern commercial pharmaceutical and biotechnology enterprise.

21.2.2 Reduced Cycle Time

Shortening the time to market is one of the major goals of a sponsor's engagement of a CRO. In today's clinical research environment, particularly involving large-scale outpatient clinical trials, the identification and recruitment of patients is usually the most significant challenge facing the trial operators. CROs applying innovations to this process, specifically designed to deal with this bottleneck, can make a major contribution to the sponsor's trial performance. For small biotechnology companies, foreign pharmaceutical operators with little development experience in the United States, and selected specialty pharmaceutical companies, engaging the full spectrum of service offerings can be very helpful. Generic companies can also benefit from selected CRO services, particularly clinical pharmacology centers for bioequivalence studies and the management of multicenter clinical trials for branded generics.

21.2.3 Accelerating the Learning Curve

Small companies with limited experience in clinical research in the targeted disease area can benefit significantly from the experience of a CRO having good performance with other product candidates in the same disease state. As mentioned above, this is particularly true in the case of small biotech startups, and can be extremely valuable to small foreign pharmaceutical and biotech operators who wish to develop their products in the United States. Such companies are interested in moving their new therapeutic entities well down the development continuum before partnering or licensing, as the value of the potential new product is greatly enhanced proportional to reduction of clinical risk.

21.2.4 Economies of Scale

Certain drug development programs can benefit from economies of scale. Multiple indications can usually be studied simultaneously at selected investigational sites. The coupling of clinical trials involving compounds targeted for depression and anxiety, osteo- and rheumatoid arthritis, and certain gastrointestinal disorders come to mind. Utilization of single investigational sites for multiple research indications provides opportunities to economize on the costs of monitoring, including travel, data collection, Institutional Review Board (IRB) activities, and electronic data capture. This approach can be effective for sponsors with limited budgets.

21.2.5 Rapid Go/No-Go Decisions

If a CRO has the capacity to rapidly initiate and successfully manage a Phase II study for a sponsor, the time to reach a milestone decision regarding the suitability of the product in patients with the targeted disease is greatly enhanced. For many years, CROs have provided opportunities for go/no-go decisions to be accelerated, simply due to the rapid start-up and initiation of the trial leading to attainment of pre-selected data points. Large pharmaceutical operators sometimes use this tactic to gain clinical information regarding safety and efficacy relative to in-licensing candidates. Such an approach can be exceptionally valuable in the area of patient safety, as Data Safety Monitoring Boards (DSMB) and IRBs will receive appropriate clinical information faster than may be possible using conventional developmental timelines.

21.2.6 Innovation

Some CROs have been slow to adopt or develop significant innovations in the clinical research process. This is primarily due to the potential threat to their revenue growth, as some of the new IT advances are specifically designed to reduce study paperwork, which is labor intensive. Developing meaningful innovations in the clinical trial process can also be expensive, and many private CROs are unable to support such investment. Public contractors often lack the sophisticated IT research expertise extant within large IT companies, and established project management procedures may be difficult to change.

Innovations that can improve the clinical trial process are the bailiwick of the niche providers, especially those in the IT space. Clinical trials can be viewed as an information transfer paradigm, and IT advances applied to companies in industries with large data loads have been immensely beneficial. Airlines, financial services, manufacturing, and the entertainment industries have all prospered from the enormous growth and effectiveness of computerized applications. The healthcare industry has been slow to adopt comprehensive IT solutions, but momentum is building, spurred in part by new government regulations requiring improved safety and lower costs throughout the healthcare delivery system.

While steady improvements in the efficiency of delivery of medical care and prescribing safety will create concomitant improvements in clinical research over time, it is likely that the pace of such advances will not satisfy the demands of sponsors for faster, higher quality and less expensive trials.

21.3 Functions of Full-Service, Multinational, Midsize, and Niche CROs

21.3.1 Full-Service, Multinational CROs

Virtually all functions necessary in medical product registration can be performed by large, multinational CROs. Projects best suited for these organizations are generally large inpatient and outpatient multicenter trials, frequently involving multiple international research sites. These organizations have the experience, financial stability, and usually the capacity to undertake large multinational trials. They can be used for freestanding services on an unbundled basis, but prefer turn-key projects which embrace the "one stop shopping" business model.

Large CROs can be used for effective planning of large clinical research projects. They commonly have extensive experience in the disease areas to be studied, although they

may not have had direct experience with some of the newer generation biologicals. A CRO may have had extensive experience with oncology product studies, but may not have been previously engaged in multicenter clinical trials involving bioengineered cancer vaccines or current generation monoclonal antibodies.

Large CROs have well-equipped, large-capacity Phase I units, often with facilities in several foreign countries. Specialized study populations are usually available. Some also have central reference laboratories allowing them to control specimen handling and analysis. Contract manufacturing is a growing revenue source for large CROs.

Several large CROs have expanded their services into related markets such as contract product detailing, patient billing services and specialty pharmaceutical product development, manufacturing, and sales. The later situation can present a conflict of interest, if their own products are potentially competitive with the sponsor product candidates. Being both a CRO and a specialty pharmaceutical firm is awkward and represents an undesirable function in the contracting business.

Additionally, these diversifications may represent sources of distraction for senior management, which can be exacerbated if the companies are publicly traded. A recent example is the diversification of the contract service organization aaiPharma (Wilmington, NC, U.S.A.) into a specialty pharmaceutical manufacturer. Quintiles (Research Triangle Park, NC, U.S.A.) also diversified aggressively in the 1990s and the acquisition of a patient billing company was problematic.

Large CROs are also expensive, thereby making them most attractive to large pharmaceutical operators, who can benefit from the increase in capacity, irrespective of cost. While the CRO market is becoming more cost-sensitive, large CROs are generally interested in large contracts, which may be a significant disadvantage to start-up biotechnology companies and medical device developers that may need cost-effective personalized service.

21.3.2 Midsize CROs

Midsize CROs frequently compete in the market on service and price. These companies may not have the financial reserves of the large CROs, but can be attractive to sponsors for projects that are regionally located, or when the sponsor has a limited budget. These companies are usually formed around specialty functions, such as data management, biostatistics, or regulatory affairs, and they add other service modalities as needed. Their experience is frequently not as extensive as the large CROs, but they can be quite helpful in selected areas, particularly where pricing issues are important. These firms can provide a more personalized service than can be expected with large CROs. These companies may concentrate on segments of the sponsor market, such as West Coast biotechnology companies.

Almost all midsize CROs are privately held, so financial information needs to be obtained through banking references, auditors, and other reliable sources. It is extremely important for sponsors to investigate the financial capabilities of these companies, as their financial stability will impact their performance and can be problematic if they are, or become, financially distressed. Careful assessment of employee turnover is also a prudent element of midsize CRO due diligence.

21.3.3 Niche CROs

Niche CROs usually specialize in selected disease areas of clinical research, or may concentrate on market segments, such as medical devices, data management, IT functions, or a myriad of other singular service offerings. Such firms can be very valuable to early stage product development enterprises, such as biotechnology start-ups, which may be focused

on one product in a single disease area. As is the case with midsize CROs, these companies require careful due diligence prior to contracting, since their financial condition and personnel stability are important.

A recent census of a large number of these small contracting groups listed more than 100 different product development functions. Many of these entities refer to themselves as CROs, but they are often small consultancies that offer a wide variety of services. Careful assessment of these service providers is necessary, as they may have a limited financial depth and operational capacity. On the plus side their fees usually are priced in the consulting range, as they have low overhead, and they may provide the customer with a high degree of expert, personalized service.

21.4 Site Management Organizations

Site management organizations provide physician investigators, qualified patients, and support personnel for inpatient and outpatient clinical trials for single site or multicenter clinical trials. These firms provide services to sponsors and CROs. They are included in the CRO industry revenue totals, and unlike CROs they incorporate per patient grant funds in their revenues.

SMOs operate with two business models; some acquire physician practices and operate them as clinical research centers exclusively, while others affiliate with existing practices and manage the administrative elements of the clinical trials. In both models, the corporate entity is responsible for marketing and sales.

Sponsors having experienced, in-house clinical development teams with attendant capacity and resources can effectively engage SMOs as part of the trials. Academic medical centers (AMCs) have functions that largely mirror those operative within commercial SMOs, especially regarding outpatient trials. As institutions, they are normally required to use their own IRB, although recently some AMCs have agreed to involvement with central IRBs.

21.5 Emerging Contract Service Providers

While not CROs in the conventional sense, emerging contract service providers must be mentioned in the context of firms providing support to sponsors in the medicinal product and device development market. These companies are large, with market capitalizations in the billions of dollars, dwarfing even the largest CROs. The prime examples of industries that participate in this arena are IT companies and central reference laboratories.

Historically, CROs have a mixed record of IT adoption; some have developed in-house custom IT solutions designed to accommodate their internal workflow and comply with their standard operating procedures (SOPs). Like sponsors, many have now acquired systems from large IT software providers, especially when such systems are employed by their clients and required by their clients for their clinical trial service providers.

The key functions of the CRO industry include improving cycle time, increasing quality and lowering the cost of therapeutic product development. A critical element of that process is the adoption of computer-based electronic enhancements (IT applied to the product development segment of the research and development continuum). Compared

to other service providers in different industries, such as financial services, transportation, and manufacturing, the CRO industry has been slow to embrace IT innovations that can significantly improve workflow and reduce costs of product development.

Before exploring the current situation and future trends regarding information technology applications in the CRO industry, it may be useful to describe some of the reasons why clinical development service providers have been "slow adopters." Described below are a variety of elements impacting CROs in their IT initiatives.

Regulators of medical products in advanced countries are generally conservative. Most regulatory agencies have physicians in positions of power and influence, and most physicians are conservative regarding unproven therapies in human subjects and patients. This conservative, risk-averse philosophy can permeate all elements of the regulatory process, including IT, where physicians frequently have little expertise. As a general observation, it is unlikely that regulators will champion IT advances in medical product development, but will look to sponsors, CROs, and IT providers to demonstrate clear utility of IT advances, especially in clinical data management and interpretation. CROs are influenced by regulatory attitudes, and may not embrace proactive philosophies regarding IT investments.

While regulatory agencies are relatively conservative, the positive impact of IT solutions on large drug approval applications containing huge volumes of clinical data is obvious. The FDA took a leadership position in encouraging electronic submissions of product approval applications in the early 1990s, with the stipulation that identical data be stored in "hardcopy" for validation of electronic data files and for potential enforcement actions. Through independent organizations such as the Drug Information Association (DIA), many workshops have been organized with significant and valuable input from the FDA. As the standards evolved, they were codified and regulations established, which reflected the FDA's position on the procedures they mandated for sponsors to comply with electronic submissions format and content. These refinements continue, and currently the activities of Clinical Data Interchange Standards Consortium (CDISC; Austin TX) represent a major leadership group for improving existing processes. Efforts of regulatory agencies and the above-mentioned independent organizations are intended to streamline the efficiency of medical product development and research.

IT advances that are currently popular in clinical research, such as content management systems, project-management software and clinical trial management software were often developed by IT providers such as Oracle, Siebel, Documentum (Pleasanton, CA), and many small clinical trial software vendors. Most of these tools offer improvements in the trial process that occurs once data have been acquired at the investigative site. The point at which data are collected from a research patient is the central locus of the site-centric process. These applications improve workflow in the "downstream" segment of the development continuum but most do not address major bottlenecks at the investigative, or "production" sites, which are "upstream." It is in these environments where the patient is evaluated by the investigator and support staff that significant improvements are needed. To increase efficiency in the "upstream" portion of this process, new IT-based solutions are necessary.

Since sponsors, including large pharma companies, are responsible for a sizable percentage of the revenues of CROs, it is understandable that their IT adoption actions would heavily impact IT choices made by CROs. In the 1990s, as IT was gaining traction in the clinical trial market, many large pharma chose to build their own proprietary IT solutions. Since they often have less robust IT capabilities than large IT companies, these efforts often failed to produce the desired results. As this trend evolved, large IT providers, especially Oracle, moved aggressively to build systems for large pharma, which capitalized on the provider's expertise and financial strength. This trend is

expected to continue, as these large IT firms provide assistance to improve many functions in medical product enterprises.

21.6 IT Companies and IT Consultants

Large IT providers such as Oracle and IBM, and others began to expand product offerings to large pharma, biotechnology companies, and CROs in the late 1980s and early 1990s. Oracle provided a software product known as Oracle Clinical in the mid-1980s, which has evolved into a larger market currently known as content management.

In the 1990s, the emergence of large consulting companies formed by the then "Big Six" accounting firms, catalyzed the rapid growth of IT in the pharmaceutical/biotechnology industry. These large consultancies, such as Accenture, Capgemini, and Bearing Point (McLean, VA), secured large contracts to design and manage enterprise-wide software for streamlining multiple corporate functions such as manufacturing, sales and marketing, finance, and research and development. These firms were drivers of IT adoption across the entire enterprise and they quickly determined that clinical research and development was an area of major opportunity for them and the IT manufacturers. As big pharma and selected biotechnology companies moved forward with software products, equipment and user-training impacting all processes under their control, CROs providing services to them were exposed to these improvements and required to procure compatible or identical systems to ensure a seamless interface in areas such as project-management, data management, and electronic submissions.

As large IT equipment manufacturers participated in the explosive growth of the market in the pharmaceutical industry, they expanded their consulting services to gain a larger percentage of this highly profitable business. IBM and others utilized this natural cross-selling opportunity to leverage package pricing to further fuel revenue growth in this attractive industry segment.

21.7 Central Reference Laboratories

Central Reference Laboratories (CRL) provide a crucial service in the product development service provider market but are not CROs in the traditional sense. Large CRLs, such as Quest Diagnostics and Laboratory Corporation of America, collect, process, and analyze human fluids and tissue samples from physician's offices, hospitals, and from their own specimen collection centers. They occupy about 33% of the total laboratory testing market, including esoteric testing involving new technologies. These include genetic assays and other biomarkers, representing the fastest-growing and most profitable segments of their businesses.

Traditionally, CRLs were early adopters of IT applications, due to the complex logistics of their large sample volumes and attendant data loads. Their specialization and growth in the procurement, transport and analysis of laboratory specimens from global clinical trial subjects and patients was largely due to the FDA's preference for central laboratory analysis of all specimens from multicenter clinical trials. Clinical trial sample analysis is a significant portion of their business, and will continue to grow as a larger number of genomic and proteomic biomarkers emerge, which will be increasingly important for determining the safety and efficacy of new medical products in the clinical testing stage.

Owing in part to the excellent, sophisticated computerized tracking systems extant within CRLs, sponsors can utilize this asset to form the basis of a simple, cost-effective, real-time trial status system. While not as potent as a Clinical Trial Management System (CTMS), this approach supports study metrics that may be adequate for multiple site trials involving nutraceuticals and natural products.

In addition to the CRLs, most large CROs have reference laboratories supporting their contracted trials. If required, they may obtain and integrate data from CRLs, as requested by the sponsor.

21.8 Benefits of Using a CRO

As discussed above, the primary role of CROs is the enhancement of economic value for sponsors in the therapeutic product development and approval process. Sponsor attitudes regarding use of CROs have evolved from extreme reluctance in 1975, to the current state of dependence. Today, the economic benefits of CRO utilization before, during, and after experimental clinical trials of new drugs, devices, diagnostics, and biologics is well-known. The principal consideration today is not whether to use a CRO, but when and which vendor. Selection of contractors is largely a function of capacity, cost, timeline, and cultural fit.

While the role of the CRO in creating economic benefit for sponsors is discussed earlier in this chapter, the relationship that is established between individuals within the firms is key to value creation. The attitudes of sponsor and vendor personnel are critical; building a dream home can be a joy for individuals who choose the right contractor for their special project. Lack of communication, poor planning, and lack of mutual respect and trust can result in undesirable outcomes. In preparing for engagement of a CRO on a clinical research project, careful planning with contingencies for unanticipated developments is absolutely essential.

A sponsor of a medical product entering the development phase must have a clinical development plan, which is created by experienced executives and consultants having expertise and experience in the disease state for which the product is intended. Once the plan is created, with appropriate input from regulators, the plan drives a clinical development budget, with consideration of the desired timeframe for projected product approval. After the plan is created, the management team determines the "make vs. buy" strategy, which determines the elements of the plan that will be outsourced.

Once identified, the contractors are selected for interviews, often in response to a Request for Proposal (RFP), at which time the sponsor and potential contractors review all aspects of the outsourced functions in the RFP. In large pharma organizations, a "gatekeeper" function acts as a filter for the vendor selection process, giving advice to product development team leaders regarding the quality, disease experience, price, reputation, timeliness, and cultural fit of the vendor with members of the sponsor's product development team.

In the negotiation phase, full-service CROs may be requested to unbundle their service offerings, although their preference is to act as "turn-key" providers for the project, offering an integrated continuum of services. The decisions regarding "turn-key" outsourcing or contracting with multiple niche providers depend largely on the outsourcing experience and expertise of the sponsor's senior management and project-management teams. This is a crucial decision, since it will impact the performance of the trial and the product approval timeframe.

Once one or more vendors have been selected, contracts are produced that detail the scope of work, timelines, budgets with milestone payment schedules, and change order

procedures. If special activities are included, such as collection and shipping of labile patient samples for genetic analysis, appropriate financial information must be included reflecting the specifications contained in the protocol. Contracts must contain provisions for early termination, whether caused by clinical holds (such as gene therapy), poor performance, regulatory irregularities, test article problems, or unforeseen issues. Audit trails of study activities, cost of supplies, and percentage of completion of patient procedures are essential, particularly if the project is cancelled or postponed.

Clinical trials involve a myriad of complex activities that need to be coordinated by motivated and experienced project managers in the sponsor and CRO organizations. Prior to the launch of a multicenter, outpatient clinical trial, qualified physician investigators, and leaders of their support staff are assembled to review the protocol, data recording methods, laboratory specimen procurement, and handling, including specialized transport requirements. Regulatory issues such as interface with external or internal IRBs monitoring by the sponsor and CRO personnel, adverse event reporting to the IRB, sponsor, and if warranted, the FDA, all must be completely understood and subject to compliance by all parties.

As indicated, a progressively larger portion of activities in a clinical trial matrix is controlled by function-specific software, provided by IT vendors, CROs, or sponsors. This may be problematic in study sites that execute multiple studies for different sponsors simultaneously, all having different SOPs, data collection procedures and payment policies. Standardized site-specific software is being developed and refined to accommodate this problem, but currently busy study sites can find this accommodation challenging.

During the conduct of the trial, constant communication between sponsor, vendor, investigator and investigation teams is critical. There is real value in this process, particularly if all parties utilize a Clinical Trial Management Software System (CTMS) that coordinates all trial activities involving the Clinical Research Coordinator (CRC) at the site and the project managers at the sponsor and CRO locations.

One of the original innovations in multicenter outpatient clinical research was the advent of Remote Clinical Data Capture (RCDC), enabling clinical sites to input visit-specific data into sponsor or CRO-specific RCDC systems. Early versions of this innovation were sponsor-friendly; and investigative sites were frequently required to use multiple systems with multiple devices, significantly increasing site labor costs, often with no additional financial support. RCDC systems have evolved during the last 20 years and are now far more effective for all parties. Today, CTMS software applications, and other electronic innovations, are creating meaningful workflow improvements.

21.9 Future Trends

Two major forces will have a monumental impact on future clinical trials: genomics and advances in information technology. Genomic advances will not only benefit basic research, by providing drug targets that can be manipulated in the laboratory but will provide new product developers and regulators with superior assessments of new therapeutic candidates. Patient populations will be available who will have precise genetic profiles, permitting clinical researchers to remove much uncertainty regarding product candidate metabolism, drug–drug interactions, toxicity, and efficacy. Several CROs have established genetic marker initiatives to discover and develop technologies that will support the next-generation trials. Patented technologies will enable them to differentiate their service offerings. In addition, they may elect to license their proprietary technologies and charge

premiums for clinical trials that will likely require far fewer patients than the current experimental designs.

The emergence of precise markers from advances in genomic and proteomic science will offer a significant competitive advantage in clinical trials for large clinical reference laboratories. As they are dependent on new diagnostic tests, which are relatively expensive, to drive growth in earnings and revenues in their conventional testing businesses, new genetic markers can be a major competitive advantage for capturing larger market share in therapeutic product development. These companies may emerge as competitors to large CROs, as they may compete to acquire selected niche service providers.

IT advances have already had a major beneficial impact on medical product development and registration. Some of the major advances in this area are electronic filing of applications for regulatory approval, RCDC, and clinical trial management software. The application of software to the analysis of the genetic profile of large patient populations is an example of the powerful intersection of pharmacogenomics and IT.

As genomic and IT advances catalyze this disruptive change, the absolute number of patients needed to document the safety and efficacy of new product candidates will decrease. This will be welcomed by sponsors who are continually pressured to lower drug and medical device costs, but it will challenge the financial performance of large CROs. Since the diversification of CROs, such as aaiPharma into specialty pharmaceuticals and Quintiles into patient billing, has been problematic, finding new growth engines will be challenging.

Among newer initiatives of large CROs is the expansion into Eastern Europe, mainland China, and India. The availability of pharmacologically naïve populations and much lower costs may lead to increased outsourcing of trials to these countries. India has a relatively educated population, well-trained clinicians, and a large number of qualified patients eligible for trials, and many are fluent in English.

The signals from the FDA regarding acceptability of trial data from these regions and countries had been unclear. CRO proposals to sponsors to conduct trials in these countries and regions may carry considerable risk. Certainly confirmatory trials, or segments of large Phase III and peri-approval studies may be acceptable to FDA, but having an entire registration based on clinical data from these areas currently carries major uncertainty. This may be especially true for advanced therapies from biotechnology companies, who can ill-afford increased risk in an already challenging regulatory environment.

References

1. Nichol, F.R., Contract clinical research: value to in-house drug development, in *Clinical Drug Trials and Tribulations*, Cato, A., Sutton, L., and Cato III, A., Eds., Marcel Dekker, New York, 2002, chap. 21.
2. *CenterWatch Directory of the Clinical Trials Industry*, Thompson Publishing, Boston, 2003.

22

The Front Lines of Clinical Research: The Industry

Lori Nesbitt

CONTENTS

22.1 Introduction

Drug development in the United States is both risky and expensive. The cost of developing a drug introduced in 1990 was approximately $500 million.[1] During the past decade, the research and development (R&D) costs of the pharmaceutical industry have doubled from approximately 12% of sales to 21%.[1] Interestingly, the average number of clinical trials required for a New Drug Application (NDA) has also doubled in the last decade to 68 per NDA.

Given the growing number of clinical trials required for FDA approval, opportunities are numerous in the provision of clinical research services. In addition, the FDA is becoming more vigilant in enforcing the ethical conduct of clinical research and the protection of research participants. Lastly, in an effort to avoid conflicts of interest or perceived improprieties, pharmaceutical and device manufacturers frequently outsource all or part of the clinical trial process to niche service providers. For these reasons, the clinical trial industry has become segmented. Each segment or service provider performs a necessary step in the clinical trial value chain. In addition to niche service providers, the growing clinical trial industry has created a need for service organizations, publications, and web sites devoted to the specialized field.

To have a thorough understanding of the industry as a whole requires a working knowledge of the freedoms, constraints, and political environments in which each service provider must operate. Perhaps a better way to stress the complexities of the industry is to realize that drug development is characterized by a high rate of failure. It is a disorderly process where very few research efforts ever bear fruit. In fact, the typical pharmaceutical company spends about 40% of its R&D budget on compounds that do not make it to market. As one expert has stated, "Drug innovation is something that is sought but not known in advance .... Only by aiming high can genuine innovation be coaxed into existence .... Innovation must be able to pay the price of failure."[1] Unfortunately, even if the new compound under study is safe and efficacious, failure can occur because of undercapitalization of the pharmaceutical company, poor data quality on the part of the sponsor or investigative site, or nonapproval by the FDA.

22.2 Clinical Trial Service Providers

Although it is estimated that 24 billion dollars was spent in human drug development (Phases I to III) in 2003, the industry is quite small and very specialized. In fact, most Americans have a limited knowledge of how new medications actually end up in their pharmacies.[1] The service providers challenged with bringing new treatments and cures to the masses include (1) Food and Drug Administration, (2) the clinical trial sponsor, (3) contract research organizations (CROs), (4) study monitors, (5) the clinical trial site, (6) site management organizations (SMOs), (7) the institutional review board (IRB), and (8) study participants.

22.2.1 Food and Drug Administration

The FDA's very noble mission is "to promote and protect public health by helping safe and effective products reach the market in a timely way, and monitoring products for continued safety after they are in use."[2] The scope of this mission has grown substantially during the past century. In addition to its consumer protection role with regard to all prescription and over-the-counter drugs, the FDA establishes and enforces standards for all food (except meat and poultry), all blood products, vaccines and tissues for transplantation, all medical equipment, all devices that emit radiation, all animal drugs and foods, and all cosmetics.

Given the complex and diverse tasks performed by the FDA, it is not surprising that there are critics on both sides of the drug approval fence. Many argue that the FDA has taken away the human right to make an informed choice regarding whether to take an investigational drug. Others argue that FDA acts too hastily in approving drugs that are later shown to be unsafe. Thus, there is a constant tension between those who want greater consumer protection and those who want greater freedom of choice.

In 1992, Congress passed the *Prescription Drug User Fee Act* (PDUFA).[3] The purpose was to establish a mechanism for financing the resources that would be needed to speed up the process of reviewing NDAs. PDUFA allows the FDA to collect user feeds from pharmaceutical companies to support the review of applications. A recent report by the U.S. General Accounting Office (GAO)[4] concludes that "PDUFA has been successful in providing FDA with the funding necessary to hire additional drug reviewers, thereby making new drugs available in the United States more quickly" (p. 6).

Supporting this conclusion is the fact that approval times have decreased from 27 to 14 months. However, the GAO also reported a small increase in the drug-withdrawal rate since the implementation of PDUFA. That is, a higher percent of approved drugs has been withdrawn from the market because of safety issues. FDA officials argue that the increase is insignificant – from 3.10% in the 8-year period before PDUFA to 3.47% in the 8-year period after.

This protection–freedom dynamic is even more intensive when it comes to drugs that are being tested for use in patients with terminal illnesses that have no other viable treatment options. In these situations, the FDA receives tremendous pressure to approve these drugs rapidly. The rationale is that the most serious risk of death from an experimental drug is no risk at all compared with the certainty of death in patients with a lethal disease.

22.2.2 Clinical Trial Sponsor

The clinical trial sponsor is an individual, company, institution, or organization that assumes responsibility for the initiation, management, or financing of the clinical trial. The sponsor is required by the FDA to conduct clinical trials to determine the safety and efficacy of the investigational agent. Safety data are usually derived through documented occurrence of adverse pharmacokinetic or pharmacodynamic effects. Alternatively, efficacy data can be evaluated by the prevention of a medical condition or through improvement of specific symptoms of a disease process.[5]

In the conduct of the clinical trial, according to the Code of Federal Regulations (CFR), the study sponsor is responsible for all aspects of the study including, but not limited to: maintaining quality assurance and quality control, medical expertise, trial design, trial management, data handling, record keeping, investigator selection, allocation of duties and functions, determining compensation to subjects and investigators, financing, notification/submission to regulatory authorities, product information, preparing and supplying study medications, monitoring and assuring that all clinical trial sites comply with federal regulatory requirements.[6]

The sponsor bears ultimate responsibility for the success, failure, and safety of the treatment under study, even after FDA approval. In addition, the sponsor is the true innovator in the clinical trial process. Innovation is expensive, causing newly available treatments to be costly to the end user. Thus, because of the escalating price of medications, the innovators are under increased scrutiny by consumers and policymakers. Paradoxically, as the population ages, the consumers are driving the demand for new cures and better treatments.

Given the high rate of "failure" in the drug industry, it is reasonable to hypothesize that drug development would take place in economies characterized by relatively free markets and prices. The ideal environment would provide adequate incentives for investing in high-risk ventures. Such an environment exists in the United States. Although the United States is home to only about 5% of the world's population, roughly 36% of the worldwide pharmaceutical R&D is conducted in the United States on a yearly basis.[1]

The United States is the world's quantitative and qualitative leader in drug development,[1] but what is the price of that leadership? Many taxpayers and consumers are outraged at the high cost of prescription drugs compared with those available in nations with price controls, such as Canada, Mexico, and the United Kingdom. *USA Today* recently featured a front-page story that compared the price of ten innovator drugs that were still under patent.[7] According to the article, the sample of drugs was 100 to 400% higher in the United States vs. Canada, Mexico, and a few European nations, where direct price controls exist. Having established this comparison, it was an easy step for the writer to conclude that what was a good deal for these other countries would be a good deal for the United States. The point that escapes the writer is that if such controls were in effect here, many of the sampled drugs would never have been developed and made available to price-controlled countries.

Citizens and policymakers misunderstand that although drug development is expensive, production costs of the pills are comparatively low. It is the formula, not the ingredients, that is costly. In addition, drug expenditure is just part of the overall expense of health care and must not be looked at in a vacuum. For example, it is estimated that, on average, U.S. citizens spend about 12% more per capita on pharmaceutical goods — or about 44 dollars per person per year more — than price-controlled nations.[1] However, a study completed by the Battelle Institute estimated that pharmaceutical research will save more than $750 billion in treatment costs for just five illnesses — Alzheimer's, AIDS, heart disease, arthritis, and cancer — during the next 25 years.[8] Being the world's leader in drug discovery is well worth the extra $44.

22.2.3 Contract Research Organizations

The daily attendance of the clinical trial process can require time, labor, and training that many sponsors feel do not match their current capabilities. Therefore, sponsors may elect to outsource any or all of their trial-related duties to a CRO. Full-service CROs offer provide data monitoring, data management, protocol development, medical writing, statistical analysis, contract management, site selection, and shipping and handling of investigational supplies. Niche CROs may elect to provide only a few of these services, such as data monitoring or medical writing. The CRO should maintain its own system of quality assurance and quality control. However, regardless of the duties assumed by the CRO, the final responsibility for the quality and integrity of the data always resides with the sponsor, and any duties not specifically transferred to a CRO remain the responsibility of the sponsor.

CROs are growing and mutating making it difficult to identify specific trends. However, one thing is certain, pharmaceutical, biotechnology, and medical devices sponsors expect to increase outsourcing to the several hundred CROs. In fact, pharmaceutical companies have already increased their use of CROs: from 28% of clinical studies in 1993 to 61% in 1999.[9] This

reflects increased spending on the CRO traditional services of Phase III study monitoring, data management, pharmacoeconomic analysis, and medical writing. Although these specialties remain the most frequently used, CROs are also offering new services to satisfy sponsors demands for faster trials and globalization.

To meet this challenge, CROs seem to be taking one of the two tracks: they are either strategically planning to become mega-CROs or niche providers. Industry observers believe the midsize CROs will disappear, mostly through merger and acquisition activity by larger CROs and by non-CROs with a strategic interest in entering the business. Although analysts forecast that within 5 years the midsize CRO will be gone, niche players with special capabilities (statistical consulting, data management, monitoring) are predicted to survive.[9] As the CRO industry consolidates, some large publicly traded CROs are making acquisitions that diversify the breadth of service beyond conducting studies. This move enables sponsors to do one-stop shopping instead of contracting with multiple companies throughout the discovery–development process. In addition, CROs are positioning themselves to gain access to populations in emerging markets such as Israel, Russia, Latin America, China, and India.

22.2.4 Study Monitors

Study monitors or clinical research associates (CRAs) can be employed by the study sponsor, CRO, or independently contracted for a specific study and, according to the International Committee on Harmonisation (ICH) and formalized by FDA in the *Guidance for Industry: Good Clinical Practice* (GCP), the purpose of a CRA is to:[10]

1. Verify that the rights and well being of human subjects are protected.
2. Verify that the reported trial data are accurate, complete, and verifiable from source documentation.
3. Verify that the conduct of the trial is in compliance with the currently approved protocol/amendment(s), with GCPs, and with applicable regulatory requirement(s).

CRAs achieve these tasks through frequent visits to the clinical trial site. During these visits, the monitor will verify source data, audit regulatory documents for accuracy and completion, perform drug accountability assessments, and communicate any concerns, problems, or new information with the study staff.

22.3 Clinical Trial Site

The front line of clinical trials is the site. It is at the site level that participants are given informed consent, study-related procedures are conducted according to the clinical trial protocol, and data are collected and reported. It is these data, aggregated from all sites, that ultimately determine the fate of the investigational drug or device. With the rigor in which clinical trials must be conducted today, site research personnel usually include the principal investigator (PI), subinvestigators, study coordinators, and regulatory managers. However, depending on the amount of research being conducted at a given location, the study coordinators are often also responsible for the regulatory compliance.

The PI is the individual who is ultimately responsible for the clinical trial at the trial site and the one who verifies that the data reported to the study sponsor are accurate. Although not required by the FDA, the PI is usually a physician. In the event that the PI

TABLE 22.1
Qualities Needed by a Successful Principal Investigator

Not Everyone Makes a Successful Principal Investigator; a Good PI
• Has an intrinsic interest in science
• Is knowledgeable about the protocol
• Always places patient care above all other priorities
• Is willing to carve out time for the study
• Is very involved in medical oversight
• Knows his or her limitations and when to ask for help
• Is tolerant of the increased need for regulatory scrutiny
• Understands that being a respected clinician does not mean being a good researcher and is open to learning about the conduct of clinical research
• Is prompt in the turnaround on documentation
• Meets participant recruitment and enrollment goals established with the sponsor

is not a physician, adequate physician oversight of the trial must be readily evident. As addressed by ICH and GCP guidance, the PI should be qualified through education, training, and experience to assume responsibility for the proper conduct of the trial and should meet all the qualifications specified by the applicable regulatory requirements (Section 4.1.1, p. 18).[10] Not all physicians are well suited for clinical research. A successful PI has distinct characteristics as listed in Table 22.1.

As defined by ICH and GCP guidelines, a subinvestigator is any individual member of the clinical trial team designated and directly supervised by the PI to perform trial-related procedures or make trial-related decisions (Section 1.56, p. 13).[10] Examples of subinvestigators include other physicians, pharmacists, pharmacologists, nurses, and study coordinators.

Clinical research coordinators (CRCs) are the research personnel who assist with patient visits, and perform study-related procedures that do not require a physician (phlebotomy, vital signs, adverse event, and concomitant medication discussions, etc.). CRCs provide the PI or physician with data required for interpretation, medical decisions (inclusion/exclusion, dosage adjustment, patient withdrawal, adverse event causality, etc.), and trial oversight. In addition, CRCs are usually responsible for transcribing source documentation (medical records, clinic notes, laboratory reports, etc.) into case report forms (CRF) supplied by the study sponsor.

Another important function of the CRC is to interact with the sponsor or CRO-appointed CRA. Because the CRA is an agent for the sponsor, the CRC–CRA relationship is one that can make or break a study. If a CRC is doing an excellent job, and the documents are available and accurate, the CRA's interactions with the site should be positive and productive. Unfortunately, this does not always happen. There are dynamics on both sides of the CRF. Some common complaints are:

- The CRA assigned to a given study changes frequently. Each CRA communicates different directives to the site, causing the site to redo work.
- The CRA has a condescending attitude toward the site and the investigator.
- The CRA is not well trained.
- The CRC is inexperienced.
- The CRA cannot obtain rapid answers to questions, often creating patient-care issues.
- The CRC cannot obtain rapid answers to queries, often extending timelines for study closure.
- The CRC makes numerous errors in the CRF.
- The CRC does not seem dedicated to the study.

In an industry where there is virtually full employment, it is difficult to find trained CRCs and CRAs. So, conflicts can arise from interaction between untrained or inexperienced personnel. Sometimes, however, personality conflicts are the main culprit. Although technology can eliminate some of the need for CRA–CRC interaction, all parties need to understand the roles and pressures on the other person.[11] For example, many CRAs travel 4 days a week and see various levels of work quality at different sites. On the other hand, CRCs are often responsible for more than one study and have requests from multiple CRAs on any given day. In addition, the CRA must respond to the needs of the research participant first, which can cause time delays in completing data queries.

Regulatory managers are usually charged with submitting regulatory documents to the IRB and study sponsor and with maintaining a regulatory binder. A regulatory binder should contain a protocol, protocol amendments, IRB approvals and correspondence, all versions of the IRB-approved patient informed consent, investigators' brochure, sponsor correspondence, curriculum vitas and licensures of the PI and subinvestigators, and any safety reports.

22.3.1 Site Management Organizations

As the number of clinical research sites have grown, a new entity, the SMOs, has arisen. SMOs, in the traditional sense, were established to offer the sponsor consolidated services at the site level. SMOs took the CRO business model and brought it to the front lines. For example, CROs offer a variety of services for the sponsor, such as site monitoring, contract administration, shipping and receiving of study supplies, and data management. SMOs offer PI recruitment, patient recruitment, and regulatory and contract management for multiple sites. As sponsors often must recruit 50 to 200 clinical trial sites, SMOs offer a one-stop shop. SMOs can provide the sponsors with multiple PIs and centralized contract and regulatory services, expediting study initiation.

SMO models vary widely in the industry. Some SMOs hire physician investigators as employees of the company. Others subcontract for investigator services. However, few offer turnkey solutions for investigators who wish to be involved in clinical research, but lack the specialized training or necessary personnel. Full-service SMOs act as a liaison between the pharmaceutical, device, or biotechnology company (or CRO), and the research patient. Services often include patient and investigator recruitment, regulatory document preparation and compliance, and study coordination.

SMOs provide an interesting entry for investigators into the clinical trial business. Specifically, some SMOs can present new investigators with clinical trial opportunities, essential training, and qualified research personnel. In turn, the investigator assumes ultimate responsibility for the ethical conduct of the study. By alleviating the physician, hospital, and health care staff from time-consuming, nonclinical tasks, SMOs can make research not only feasible but also lucrative for investigators and hospitals. This risk-sharing model can be beneficial to all parties.

22.3.2 Institutional Review Boards

Since the *Kefauver–Harris Amendment*[12] was adopted in 1962, pharmaceutical manufacturers have been held responsible by the FDA for providing new medications that are both safe and effective. In addition to the *Kefauver–Harris Amendment*, the *Belmont Report* (written in 1979),[13] established the ethical principles and guidelines for conducting research. The FDA requires that clinical trials be conducted in compliance with a protocol that has been

approved by an IRB or Independent Ethics Committee (IEC). The terms IRB or IEC refer to any board, committee, or other group formally designated by an institution to review, approve initiation, and conduct periodic review of research involving human participants.[14] The FDA expects the IRB to review all research-related documents and activities that pertain directly to the rights and welfare of the participants of proposed research. The IRB has the authority to approve, require modification, or disapprove all research activities as specified in the federal regulations. The primary purpose of the IRB and this formal review process is to protect the rights and welfare of human participants involved in these clinical trials. It is the federally mandated charge of the IRB to ensure the safety of the research participant.

However, the FDA is not the only agency governing the function of an IRB. The FDA only oversees clinical trials when they involve an FDA-regulated product (i.e., drug, device, or biologic). The Department of Health and Human Services (HHS), Office for Human Research Protection (OHRP) oversees federally funded research. The *Federal Policy for the Protection of Human Subjects*,[15] known as the "Common Rule," is the basic HHS policy for protection of human subjects, now codified in the CFR.[16] Even though the FDA regulations are not part of the "Common Rule," the basic requirements for IRB's and informed consent are congruent.

The differences between the "Common Rule" and the FDA regulations are differences in applicability. HHS regulations are based on federal funding, and FDA regulations are based on the use of FDA-regulated products. Examples of some of the differences in the FDA regulations and the "Common Rule" include the following:

- Differing definitions of a "human subject."
- HHS discusses "research" whereas FDA discusses "clinical investigations."
- FDA makes no provisions for waiving informed consent whereas HHS regulations provide for certain conditions in which an IRB can waive or alter elements of informed consent.
- FDA regulations states that subjects must be informed about FDA inspections.
- FDA requires that subjects sign and date the informed consent.
- FDA allows the sponsor to request a waiver of IRB review; under HHS regulations, certain categories of research are exempt, and department heads can waive regulations.
- HHS and FDA both have adopted additional protections for children, but HHS also has additional protections for fetuses, pregnant women, and prisoners.
- FDA has responsibility and authority over all parties participating in FDA-regulated research and has regulations unique to product review responsibilities.

IRBs should be knowledgeable in both FDA and HHS regulations and should know which regulations to apply when conducting its review. FDA regulations apply when products are regulated by the FDA. The "Common Rule" applies when the research is federally supported or conducted or when it is being conducted in an institution that has agreed to review all research under the "Common Rule." Both rules apply if the research is federally funded and involves FDA-regulated products or if the FDA-regulated research is being conducted in an institution that has agreed to review all research under the "Common Rule." As a rule of thumb, no matter which regulations apply, IRB members should use the *Belmont Report*[13] and its guidelines in all of their daily decisions.

22.4 The Evolution of Ethical Principles in Clinical Research

As the demand to develop new and improved medical treatments has grown so has the need for human participants to test the potential new treatments. The federal government has taken an increasingly prominent role in the oversight of human subject protection. The *Belmont Report*[13] established the ethical principles and guidelines for conducting research. It describes the three, basic, ethical principles (respect for persons, beneficence, and justice) that are relevant to clinical research involving humans. These principles dictate to the clinical research industry that it does not have the right to use people for scientific benefit without their permission — no matter how noble the cause may seem. The research must do something good, and it has to be fair — people cannot be exploited for research purposes. Applying these ethical principles brings the research industry closer to providing sound, ethical research. The regulations that govern the conduct and oversight of clinical research involving human participants were written to implement these ethical principals.

The history of the development of formal ethical principles of research with humans can be traced to World War II. Physicians in Nazi Germany performed "medical experiments" on thousands of concentration camp prisoners. These experiments included determining how long humans would survive in freezing water and high altitudes, injecting people with viruses, and forcing people to ingest poison. Whereas the German physicians argued that the experiments were medically justified, the Nuremberg Military Tribunal declared them to be "crimes against humanity." Seven of the physicians were sentenced to death. The judges who wrote the verdict in 1947 included a section on medical experiments, which became known as the *Nuremberg Code*.[17] This code was the first formal set of ethical principles for researchers. The first item in the *Nuremberg Code* emphasizes obtaining informed consent as a primary responsibility of the researcher:

> The voluntary consent of the human subject is essential. This means that the person involved should have legal capacity to give consent; should be so situated as to be able to exercise free power of choice, without the intervention of any element of force, fraud, deceit, duress, over-reaching, or other ulterior form of constraint or coercion; and should have sufficient knowledge and comprehension of the elements of the subject matter involved as to enable him to make an understanding and enlightened decision. This latter element requires that, before the acceptance of an affirmative decision by the experimental subject, there should be made known to him the nature, duration, and purpose of the experiment; the method and means by which it is to be conducted; all inconveniences and hazards reasonably to be expected; and the effects upon his health or person, which may possibly come from his participation in the experiment.
>
> The duty and responsibility for ascertaining the quality of the consent rests upon each individual who initiates, directs, or engages in the experiment. It is a personal duty and responsibility, which may not be delegated to another with impunity.[17]

In 1964, the World Medical Association adopted the "Declaration of Helsinki"[18] to guide researchers in the ethical conduct of medical research. This document has been revised several times since originally adopted. These guidelines reinforced the importance of informed consent, although it added a provision to allow for that consent process to be performed with a legal guardian for people unable to provide consent for themselves. The Declaration of Helsinki also delineated the difference between clinical research, defined as "medical research in which the aim is essentially diagnostic or therapeutic for a patient" and nonclinical biomedical research, defined as "medical research, the essential object of which is purely scientific and without implying direct diagnostic or therapeutic value to the person subjected to the research."

At the time of unfolding this history of medical research ethics internationally, a heinous medical experiment was being conducted in the United States in Tuskegee by the U.S. Public Health Service. The purpose of this research, which began in the 1930s, was to study the natural history of untreated syphilis. More than 400 African Americans with syphilis were recruited into the study without their permission. They were also deceived; they were told that some study tests were "special free treatment." During the course of this study, the researchers learned that the mortality rate was twice as high in subjects with syphilis when compared with controls without syphilis. In the 1940s, penicillin was proven an effective treatment for syphilis. Despite this finding, the research continued. The subjects were not told of the available treatment, and it was not given to them.

Reports of the Tuskegee study began appearing in the media in 1972. Overwhelming and justified public shock and anger resulted in federal action to compensate for the egregious ethical breaches and to prevent them from recurring. The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research was established in 1974 to identify the basic ethical principles to guide all research with humans. The Commission's work was documented in *The Belmont Report — Ethical Principles and Guidelines for the Protection of Human Subjects*.[13] That report contained the three ethical principles mentioned above that guide the work of modern researchers and that became the foundation for federal regulations governing research with humans: respect for persons, beneficence, and justice. The principles are described in greater detail below.

22.4.1 Respect for Persons

The first of the ethical principles reflects the emphasis in the Nuremberg Code on informed consent. That is, respect for research participants means treating them as autonomous agents. The *Belmont Report* defines "autonomous agent" as "an individual capable of deliberation about personal goals and of acting under such deliberation ...." This means that the informed consent process is needed so that the potential participants are given all the information they need to determine whether participating in the study is in their best interest. The principle of respect also means that researchers conducting trials with "vulnerable" populations (e.g., children or the cognitively impaired) should ensure that extra provisions are made to protect those participants who have diminished autonomy. Of course, the principle of respect includes the notion that the consent process must be completed with no coercion or pressure by the researchers to participate.

22.4.2 Beneficence

The *Belmont Report*'s principle of beneficence refers to the need to ensure that all aspects of a study are designed to obtain the desired knowledge in a way that maximizes benefits and minimizes risks to the participants. The principle of beneficence also means that a risk–benefit analysis must be performed on every proposed study. In determining if this ratio is ethical, consideration must be given to the impact on both the participants and society.

22.4.3 Justice

The principle of justice refers to the fairness or equity involved in the selection of research participants. The point is to ensure that risks or benefits incurred by research participants are not unfairly concentrated in one segment of the population.

22.4.4 How the Investigator Applies Ethical Principles

An investigator applies these ethical principles daily by ensuring that the informed consent process is performed perfectly with every research participant. This means that all staff involved in the consent process must be well trained, not only in the research protocol in question, but also in the proper elements of informed consent. A good rule of thumb for the people conducting consent discussions is for them to think about what they would want their mother, father, spouse, or friend to know about the study to enable them to make an informed decision. The investigator should never allow this process to be rushed. In addition, to reduce the possibility that the potential participant feels pressured by the power differential that often characterizes the doctor–patient relationship, some ethicists argue that the physician should not be the primary person conducting the consent discussion.

The researcher complies with the concept of beneficence in two ways. First, the investigator carefully reviews the study protocol to determine whether the design provides a generally favorable ratio between the potential risks to the subjects and the potential benefits to the subject and society. If the researcher decides to conduct the study, the second strategy is to review carefully each potential candidate to ensure that this ratio is favorable for each potential participant. Of course, the investigator also encourages the potential participant to weigh the risks and benefits for themselves through the consent process.

Finally, the investigator fulfills the principle of justice by ensuring that his or her subject-recruitment strategies do not systematically discriminate against any specific group. In the Tuskegee study, all of the risks were concentrated among African American males. In other cases, the benefits might be concentrated in a particular group. One of the current ethical dilemmas facing research professionals is the conduct of research involving children. Because of the focus on protecting vulnerable populations, drug research in children has been very limited. Thus, the benefits of research were concentrated in adults. The result was that physicians had little scientific basis for prescribing medications in children. To improve the practice of pediatric medicine and to expand the benefits of research beyond adults, the federal government established incentives to pharmaceutical companies to sponsor research with children. This, in turn, led to additional regulations designed to protect the rights of minors who participate in research. This is just one example of the delicate balance of justice in selecting research populations.

22.5 Regulations Governing Institutional Review Boards

Investigators embarking on research path might be tempted to focus their energies exclusively on those aspects of the study that involve direct patient care and medical decisions. They might be tempted to skip the remaining portions of this chapter that focus on the regulations governing IRBs. Why, after all, does an investigator need to know about IRBs?

The answer is that the FDA holds investigators responsible for ensuring that the studies they conduct are reviewed by an IRB that meets FDA regulations governing the function of IRBs. More specifically, when an investigator signs FDA Form 1572,[19] the investigator is making the following commitment (among others): "I will ensure that an IRB that complies with the requirements of 21 CFR Part 56 will be responsible for the initial and continuing review and approval of the clinical investigation." Therefore, it is prudent for the investigator to take care in selecting an IRB.

22.5.1 What Institutional Review Board Members Should Know about Clinical Research

Although IRB members can and should come from a variety of backgrounds, all board members should know the basic elements of clinical research and should stay focused on their role in the ethical review process. Any board member who does not have a research background should take the time to learn the basics of the industry. Books, such as this one, can introduce the language, regulations, and ethical issues in the conduct of research with humans.

Board members should look carefully at an investigator's qualifications for conducting the research. If a protocol proposes to study the effects of a new treatment for hypertension, the investigator should not be a podiatrist. The investigator is responsible for overseeing the medical care and conduct of the trial and, therefore, must have the proper qualifications to accomplish that task.

Board members should understand the need to avoid conflicts of interest. Any board member with a conflicting interest in a study should not participate in the deliberations or voting on that study. The IRB may ask that member to leave the room during deliberations and voting to avoid a political environment that is not conducive to objective ethical review. However, it is the IRB's decision whether an individual should remain in the room.

Finally, board members should understand that clinical research involves tremendous coordination of systems, staff, and supplies; however, it is *not* the responsibility of the IRB to plan, facilitate, or verify that the investigator has resolved the logistics of a given study. The logistics are something the investigator works out with the sponsor. Instead, the IRB should stay focused on evaluating the risks and benefits of conducting the study.

22.5.2 What Institutional Review Board Members Should Know about Their Responsibilities

IRB members should be aware that their duties include (but are not be limited to) evaluating proposed investigations and approving or disapproving the investigation after considering the medical soundness in light of the rights and safety of the human participants involved; determining compliance with acceptable standards of professional conduct and practice; and assessing community, ethical, and moral values. The members also evaluate the qualifications of the PI with emphasis placed on that individual's professional development as it relates to the degree of protocol complexity and the risk to human research participants. Because the primary purpose of an IRB is to ensure – in advance and by periodic review – that appropriate measures are taken to safeguard the rights, safety, and well-being of human participants involved in a clinical trial, the board members must determine that the following requirements are satisfied:

- Human participants are protected from ill-advised research or research protocols in light of both ethical and scientific concerns.
- Risks to participants are minimized.
- Risks to participants are reasonable in relation to anticipated benefits and the importance of the knowledge that may be expected to result.
- Selection of participants is equitable.
- Informed consent will be obtained from each prospective participant or the participant's legally authorized representative and will be documented in accordance with IRB, FDA, and ICH informed consent regulations and guidelines.
- Adequate provisions are made for monitoring the data collected to ensure the safety of participants.

- Adequate provisions are made to protect the privacy of participants and maintain confidentiality of data.
- Appropriate additional safeguards have been included to protect the rights and welfare of participants who are members of a particularly vulnerable group,[14] such as persons with acute or severe physical or mental illness or persons who are economically or educationally disadvantaged.
- Participant selection and exclusion criteria — including justification of the use of special participant populations, such as children, pregnant women, human fetuses, and neonates or the mentally handicapped — are appropriately established.
- The study design includes a discussion of the appropriateness of the research methods.

In addition to the above, the IRB has additional and special review requirements to protect the well-being of children when children are to be involved in the research (Subpart D, Sections 401–409).[16]

22.6 Risks-vs.-Benefits Analysis: The Human Advocate

To ensure that the rights and welfare of human research participants are protected, the IRB takes on the task of performing a risk vs. benefit analysis. Before a clinical trial can be initiated, foreseeable risks and inconveniences should be weighed against the anticipated benefit for the trial participant and for society (Section 22.2).[10] Risks associated with participation in research should be justified by the anticipated benefits; only then, can the trial be initiated. Because this requirement is clearly stated in the federal regulations, it is, therefore, one of the major responsibilities of the IRB to assess the risks and benefits associated with proposed research. Definitions of the terms that the IRB uses to assess risk include the following:

- *Benefit* — a valued or desired outcome, an advantage.
- *Minimal risk* — the probability and magnitude of harm or discomfort anticipated in the proposed research are not greater, in and of themselves, than those ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests.
- *Risk* — the probability of harm or injury (physical, psychological, social, or economic) occurring because of participation in a research study. Both the probability and magnitude of possible harm can vary from minimal to significant.

The IRB must determine whether the anticipated benefit – either of new knowledge or of improved health for a research participant – justifies asking individuals to expose themselves to the potential risks. The IRB's assessment of risks and anticipated benefits involves the following steps:

- Identifying the risks associated with the research in comparison to the risks of treatments the participants would receive if not participating in research.
- Determining that the risks will be minimized to the extent possible.
- Identifying the probable benefits to be derived from the research.
- Determining that the risks are reasonable in relation to the benefits to the participants, if any, and the importance of the knowledge to be gained.

- Assuring that potential participants will be provided with an accurate and fair description of the risks or discomforts and the anticipated benefits.
- Determining intervals of periodic review and, where appropriate, determining that adequate provisions are in place for monitoring the data collected.

In addition, the IRB should determine that the provisions to protect the privacy of participants and to maintain the confidentiality of the data are adequate. When participants are members of a vulnerable population, the IRB should ensure that appropriate additional safeguards are being used to protect the rights and welfare of these participants.[20]

22.6.1 Identifying and Assessing the Risks

In the process of identifying what constitutes a risk, only risks that can result directly from the research should be considered, but before an activity involving participants is eliminated from the risk–benefit analysis as therapy, the IRB should be very sure that the activity is not actually research. IRB member should be aware that the potential risks faced by research participants could be associated with the design mechanisms used to ensure valid results and by the interventions that may be performed during the course of the research. Designs involving randomization to treatment groups have the risk that the participant may not receive a treatment that could be effective. Participants involved in a double-blinded study are at risk of the necessary information for individual treatment not being available to the proper individuals when needed. The added risk of invasion of privacy and violations of confidentiality are also possible within the methods used for gathering information in behavioral, social, and biomedical research.

Risks to which research participants may be exposed have been grouped as physical, psychological, social, and economic harms.[21] The description of each classification is as follows:

- *Physical harms*: Medical research often involves exposure to minor pain, discomfort, injury from invasive medical procedures, or harm from possible side effects of drugs. All of these should be considered "risks" for purposes of the IRB review.
- *Psychological harms*: Participation in research may result in undesired changes in the thought processes and emotions. These changes may be temporary, recurrent, or permanent. IRB members should be aware that some research has the potential to cause serious psychological harm. Stress and feelings of guilt or embarrassment may arise from thinking or talking about one's own behavior or attitudes on sensitive topics. These feelings may occur when a participant is being interviewed or filling out a questionnaire. IRBs will confront the possibility of psychological harm when reviewing behavioral research that involves a component of deception, particularly if the deception includes false feedback to the participants about their own performance.
- *Invasion of privacy*: Access to a person's body or behavior without consent constitutes an invasion of privacy. The IRB must determine if the invasion of privacy is acceptable in light of the participant's reasonable expectations of privacy in the situation under which the study is being performed and if the research questions are of sufficient importance to justify this intrusion. The IRB should determine if the research could be modified so that the study could be conducted without the invasion of the participants' privacy.

- *Breach of confidentiality*: Confidentiality of data requires safeguarding information that has been given voluntarily by one person to another. Some research requires access to the participants' hospital, school, or employment records. Such access is generally acceptable as long as the researcher protects the confidentiality of that information. The IRB should be aware that a breach of confidentiality could result in psychological or social harm.
- *Social and economic harms*: Some invasions of privacy or breaches of confidentiality could result in embarrassment with one's business or social group, loss of employment, or criminal prosecution. Confidential safeguards must be strong in these instances. Examples of these particular sensitivities include information about alcohol or drug abuse, mental illness, illegal activities, and sexual behavior. Participation in research may also result in additional costs to the participant.

22.6.2 Minimal Risk vs. Greater-than-Minimal Risk

Once the risks have been determined, the IRB must then assess whether the research involves greater-than-minimal risk. Regulations governing the functions of an IRB allow approval through the expedited review process for research projects that contain no more-than-minimal risk or that involve participants only in one or more approved categories.

In research involving more-than-minimal risk, potential participants must be informed of the availability of medical treatment and compensation for a research-related injury, including who will pay for the treatment and the availability of other financial compensation.[22] Institutions are not required to provide care or payment for research injuries; however, some institutions provide hospitalization and necessary medical treatment in an emergency situation.

22.6.3 Vulnerable Populations and Minimal Risk

When research involves especially vulnerable populations (e.g., fetuses and pregnant women, prisoners, children, mentally disabled) regulations strictly limit research involving more-than-minimal risk. Special limitations are recommended when the research involves individuals who are institutionalized or mentally disabled. For these situations, it is recommended that minimal risk be defined in terms of the risks normally encountered in the daily lives or the routine medical and psychological examination of healthy participants. In these cases, the IRB should determine whether the proposed participant populations would be more sensitive or vulnerable to the risks involved by the research because of their general condition or disabilities. These concerns are also equally applicable to other participants (e.g., taking blood samples from a hemophiliac, outdoor exercises with asthmatics if the air is polluted, changes in diet for a diabetic, and giving over-the-counter drugs for minor ailments to pregnant women).

22.6.4 Determining Whether Risks Are Minimized

The IRB is responsible for assuring that risks to participants are minimized. In assuring that the risks are minimized, the IRB should obtain and review the protocol, including the investigational design, scientific rationale, and the statistical reason for the structure of the proposed research. Results from previous studies (i.e., the investigator's brochure) should also be reviewed during this process. The expected beneficial and harmful effects within the research, as well as the effects of any treatments that may be ordinarily administered, and those associated with receiving no treatment should also be analyzed.

Whether potential harmful effects can be detected, prevented, or treated should be considered. Risks and complications of any underlying disease that may be present should be assessed as well.

The IRB should determine whether or not the investigators are competent in the area of the proposed research and whether they are serving in more than one role, which may complicate their interactions with participants. Potential conflict-of-interest issues should be identified and resolved before IRB approval.

Deciding whether the research design will produce useful data will assist the IRB in determining whether risks are minimized. Participants may be exposed to risk without sufficient justification when the research design does not contain a sample size large enough to produce valid data or conclusions. Faulty or poor research design means that the risks are not likely to be reasonable in relation to the benefits. Sometimes, procedures that are included for purposes of good research design, but which add disproportionate risks to participants, may be unacceptable. Assuring that adequate safeguards are incorporated into the research design is a useful method of minimizing the risks.

22.6.5 Assessing Anticipated Benefits

Benefits of research are considered benefits to participants and to society. Research often involves the evaluation of procedures that may benefit the participants by improving their conditions or providing a better understanding of their diseases or disorders. In this type of research, participants undergo treatment for a particular illness or abnormal condition. Patients and healthy volunteers may choose to take part in research that is neither related to any illness or condition they might have nor structured to provide any diagnostic or therapeutic benefit. This type of research is designed to gain knowledge about human behavior and physiology. Research that contains no immediate therapeutic intent may benefit society overall by providing increased knowledge, improved safety, advances in technology, and overall better health. Anticipated benefits to the participant and the expected knowledge to be gained should be clearly stated within the protocol.

22.6.6 Determining Whether Risks Are Reasonable

The IRB must consider a number of factors when determining of whether the risks are reasonable in relation to the anticipated benefits. The evaluation of the risk–benefit ratio is an ethical judgment that an IRB must make, and each case must be reviewed separately. This judgment often depends on subjective determinations and community standards. When making its decision, the IRB relies on currently available information regarding the risks and benefits of the interventions from previous animal and human studies (from the investigator's brochure), and the extent of confidence in the knowledge. However, human responses may differ from those of animals; therefore, although that information may suggest possible risks and benefits to humans, it is not conclusive. Within its assessment, the IRB should also consider the proposed participants and be sensitive to the different feelings and views individuals may have about risks and benefits.

The risk–benefit assessments depend on whether the research involves the use of interventions that have the intent and reasonable possibility of providing a benefit to the participant, or whether it only involves procedures for research purposes. In research containing interventions expected to provide direct benefit to the participants, a certain amount of risk is justifiable. In research where no direct benefits are anticipated, the IRB should evaluate whether the risks presented by procedures only to obtain generalized knowledge are ethically acceptable.[20]

22.6.7 Continuing Review and Monitoring of Data

Regulations governing the functions of an IRB require that an IRB reevaluate research projects at intervals appropriate to the degree of risk but not less than annually.[16] The reevaluation is performed to review the entire research project again and to reassess the risk–benefit ratio, which may have changed since the last review. During the course of a study, new information regarding the risks and benefits, unexpected side effects, unanticipated findings involving risks to participants, or knowledge resulting from another research project may become apparent. The IRB should determine whether these situations have occurred and whether there is any additional information regarding risks or benefits that should be revealed to the participants.

The investigator should be aware that the interval for IRB review can be more frequent than once per year. If an investigator allows an IRB approval date to lapse, the research no longer has IRB approval. No approval means that the research must stop. Extensions of approvals do not exist. Investigators should provide the IRB with an adequate detailed report in a timely manner to avoid the approval of their study from expiring.

22.7 Informed Consent: Not Just a Document

One of the most important pieces of a research trial is the informed consent of the research participant. However, a signature on an informed consent document does not constitute the end of the informed consent. It is very important, for the research community (the sponsor, the investigator, and the IRB), to remember that informed consent is not just a document, but is, instead, a continuing process that carries through to the end of the individual's participation in the trial.

The informed consent process begins with the recruitment of potential participants, whether by radio advertisement, flyer, or initial contact with the PI or the research staff. The initial consent involves explaining to a potential participant the risks, benefits, alternatives, procedures, and purpose of the study; allowing that individual to ask questions and providing satisfactory answers to those questions; sharing new information with the participants as it becomes available during the course of the study because it may change their willingness to continue participation; and for long-term studies, revisiting the consent because of possible capacity changes.

The written informed consent form should be presented to potential participants in a language that they understand and written in terms that they can comprehend. This is one of the important functions of an IRB. The IRB members review every informed consent form to determine if it contains all of the required elements and any additional required elements of an informed consent form as set forth in the governing regulations. The informed consent form is also reviewed to determine that complete, accurate, and pertinent study-related information is being provided to the potential participants and that medical terms are clearly defined, in simple language that the study population can understand.

Human participant protection is a shared responsibility between the sponsor, the investigator, and the IRB. It is their responsibility as a team to ensure that the participants remain well informed and that their rights and welfare are protected. It is important that all members of the research team understand the informed consent form regulations,[23] their part of the informed consent process, and apply this knowledge to each informed consent form that is reviewed, presented, and distributed to a potential participant.

22.7.1 Elements of an Informed Consent

The investigator is responsible for providing every potential research participant with complete, accurate, and pertinent study-related information while adhering to the applicable governing rules and regulations. The process of providing this information to potential participants is informed consent. Except as provided in the regulations,[24] no investigator may involve a human being as subject in research unless the investigator has obtained the legally effective informed consent for the participant or the participant's legally authorized representative.[25] The exceptions include the participant being in a life-threatening situation that makes it necessary to use the test article; being unable to communicate with the participant to obtain a legally effective consent; insufficient time to obtain the consent from the participant's legally authorized representative; and no available alternative method of approved or generally recognized therapy that provides an equal or greater chance of saving the participants life.[26]

The governing rules and regulations require certain elements (information) be included within every informed consent form that is provided to a potential research participant. These required elements are outlined in Table 22.2.

22.7.2 Additional Required Elements

When appropriate, one or more of the following elements of information should also be provided to each participant. Note that additional required elements are not optional but are required when applicable:

- A statement that the particular treatment or procedure may involve risk to the participant (or to the embryo or fetus if the participant is, or may become, pregnant and to nursing infants) that is currently unforeseeable.
- Anticipated circumstances under which the participant's participation may be terminated by the investigator without regard to the participant's consent.
- Any additional costs to the participant that may result from participation in the research.
- The consequences of a participant's decision to withdraw from the research and procedures for orderly termination of participation by the participant.
- A statement that significant new findings developed during the course of the research, which may relate to the participant's willingness to continue participation, will be provided to the participant.[26]

In addition to the required elements of an informed consent form, an IRB can and may require other standard information or signatures be added to all informed consent forms being reviewed by the board.

22.8 What the Site Should Know about Institutional Review Boards

When choosing an IRB, the investigator may have the option of using a "local" IRB or a "central" IRB. A local IRB is one that is housed within an institution and has been developed for overseeing research conducted within the institution or by the staff of the institution. For example, many community hospitals and academic medical centers have their own internal IRB, composed largely of clinicians who conduct research at that institution.

TABLE 22.2

Elements Required in an Informed Consent

- A statement that the study involves research
- An explanation of the purpose of the research and the expected duration of the participant's participation
- A description of procedures to be followed and identification of any procedures that are experimental, including all invasive procedures
- The approximate number of participants involved in the trial and the participant's responsibilities
- A description of any reasonably foreseeable risks or discomforts to the participant
- A description of any benefits to the participant or to others that may reasonably be expected from the research; when there is no intended clinical benefit to the participant, the participant should be made aware of this fact
- A disclosure of appropriate alternative procedures or courses of treatment that might be advantageous to the participant and their important potential benefits and risks
- A statement describing the extent of confidentiality of records, which notes the possibility that the monitors, auditors, IRB, and the regulatory authorities may inspect the records for verification of clinical trial procedures and data without violating the confidentiality of the participant to the extent permitted by the applicable laws and regulations and that, by signing a written informed consent form, the participant or the participant's legally acceptable representative is authorizing such access; that records identifying the participant will be kept confidential and, to the extent permitted by the applicable laws or regulations, will be made publicly available; and that if the results of the trial are published, the participant's identity will remain confidential
- An explanation of whether compensation or medical treatment is available if injury occurs and, if so, what they consist of, or where further information may be obtained
- The anticipated prorated payment, if any, to the participant for participating in the trial
- A statement of dosage/frequency and the probability for random assignment to each treatment
- An explanation of whom to contact for answers to pertinent questions about the research and research participants' rights and whom to contact in the event of a research-related injury to the participant
- A statement that participation is voluntary, that refusal to participate will involve no penalty or loss of benefits to which the participant is otherwise entitled, and that the participant may discontinue participation at any time without penalty or loss of benefits to which the participant is otherwise entitled
- A place for signatures by the participant, physician, person obtaining consent, and witness (less signatures may be required); it depends on the policies and procedures of the particular IRB

A central IRB is an "outside" or "independent" IRB. The central IRB serves to review research for non-IRB institutions, private practices, or outpatient clinical trials. Investigators conducting research in a noninstitutional setting often choose to use an established central IRB rather than forming their own. In addition, many hospitals and other institutions are choosing to eliminate their internal IRBs and outsource this function. The benefits of the decision to outsource include:

- Reduced likelihood of conflict of interest or the appearance of bias because the board members reviewing the proposed protocols are unlikely to be friends, colleagues, or acquaintances of the researchers whose work they are reviewing.
- Reduced liability by selecting to outsource to an organization whose sole focus and expertise is in ensuring compliance with the many regulations governing the work of IRBs.
- Potential cost savings by eliminating staff currently dedicated to performing IRB duties by outsourcing to an IRB that achieves economies of scale through higher volumes of reviews.

An investigator must be qualified by education, training, and experience to assume responsibility for the proper conduct of a clinical trial and should meet all qualifications specified by the applicable regulatory requirements. Investigators must agree to abide by the decisions of their selected IRB and should comply with all governing rules and regulations.

22.8.1 What Research Requires Institutional Review Board Approval and Oversight

According to the *Belmont Report*,[13] items that must go before an IRB for review, for the protection of human participants, include any projects that include any element of research in an activity. The definition of research is described as a systemic investigation designed to develop or contribute to generalized knowledge.[16] Medical research involving human participants also includes research on identifiable human material and identifiable data.[18] The obvious material requiring submission to, and review by, an IRB would be industry-funded research, which would include biomedical, behavioral, medical devices, and humanitarian-use devices. If there is confusion about whether or not a project is research, several questions can be asked to help make this determination:

- Is there a hypothesis?
- Does it include research development?
- Will the knowledge be used outside the institution?
- Is there a question to be answered for reasons other than clinical care or routine evaluation?
- Is there an intent to generalize the information?
- Is there a specific intent?
- Is the purpose for the scientific community?
- If it were not going to be published, would you do it anyway?

Investigators should be aware that some research projects may qualify for expedited review. Research that would qualify for this type of review includes certain categories of research that involve no more-than-minimal risk. A list of research categories that qualify was published in the *Federal Register* in 1981 (Table 22.3).

TABLE 22.3

Categories of Clinical Investigations That Can Go Through an Expedited Review[28]

1. Collection of hair and nail clippings in a nondisfiguring manner, of deciduous teeth, and of permanent teeth if patient care indicates a need for extraction
2. Collection of excreta and external secretions including sweat and uncannulated saliva, of placenta at delivery, and of amniotic fluid at the time of rupture of the membrane before or during labor
3. Recording of data from subjects who are 18 years of age of older using noninvasive procedures routinely employed in clinical practice. This category includes the use of physical sensors that are applied either to the surface of the body or at a distance and do not involve input of matter or significant amounts of energy into the subject or an invasion of the subject's privacy. It also includes such procedures as weighting, electrocardiography, electroencephalography, thermography, detection of naturally occurring radioactivity, diagnostic echography, and electroretinography. This category does not include exposure to electromagnetic radiation outside the visible range (e.g., x-rays or microwaves)
4. Collection of blood samples by venipuncture, in amounts not exceeding 450 ml in an 8-week period and no more often than two times per week, from subjects who are 18 years of age or older and who are in good health and not pregnant
5. Collection of both supra- and subgingival dental plaque and calculus, provided the procedure is not more invasive than routine prophylactic scaling of the teeth and the process is accomplished in accordance with accepted prophylactic techniques
6. Voice recordings made for research purposes such as investigations of speech defects
7. Moderate exercise by healthy volunteers
8. The study of existing data, documents, records, pathological specimens, or diagnostic specimens
9. Research on drugs or devices for which an investigational new drug exemption or an investigational device exemption is not required

22.8.2 Requirements of an Investigator for Institutional Review Board Research Approval

For an investigator to obtain IRB approval for a research proposal, required information should be provided to the prospective IRB for review and consideration in a timely manner. PIs should know what their IRBs expect, and should ask questions early in the process to avoid any delays. Investigators should remember that an IRB review is based on human-participant concerns, and should expect questions. It is very helpful for the investigator to contact the IRB before making a submission to inquire about its document-submission requirements. Incomplete submissions to the IRB will definitely cause the review of the research project to be delayed. When IRBs have to request a number of modifications or seek additional information, IRB approval must be deferred, pending subsequent review by the full board upon receipt of the requested information. Providing the IRB with all required documents with the initial submission, however, will help eliminate undue delays and make the IRB review process more expedient. The pertinent information that the investigator will need to provide may vary from IRB to IRB; however, in general, the standard, required information would include the material listed in Table 22.4.

TABLE 22.4
Information that IRBs May Be Required from a Principal Investigator (PI)

Standard Information Required
Brief letter, memorandum, or note requesting approval of the research
Completed, signed original of the IRB application or Information and Site Survey (if required by the IRB)
Investigational Device Exemption (IDE), Humanitarian Device Exemption (HDE), or Premarket Approval (PMA) numbers for device trials
Copy of the protocol/amendments; (title of study; purpose, including expected benefits obtained by doing the study; participant selection criteria; par ticipant exclusion criteria; study design, including as needed, a discussion of the appropriateness of research methods; description of methods performed)
Written informed consent document and consent form updates (which include all of the required elements of an informed consent — asee Section 22.6)
Completed FDA 1572 form, if applicable (not required for device trials)
Investigator's brochure (if applicable)
Curriculum vitae and licensure for the principal investigator and each subinvestigator
Disclosure of any payments or compensation to a participant for participating in the study
Name of the sponsor
Results of previous related research (i.e., investigator's brochure)
Participant recruitment procedures, materials, or advertisements
Written information that will be provided to the participant
Provisions for managing adverse reactions
Justification for use of special/vulnerable participant populations; as well as additional safeguards that will be used to protect these participants (i.e., the mentally retarded, children, prisoners, pregnant women, etc.)
A disclosure of any compensation provided to the participant for participating in the study
A disclosure of extra costs to the participant because of participation in the study
Protection of the participant's privacy
Additional Items the PI Should Provide
Any changes in the study after initiation
A report of any unexpected serious adverse reactions or information regarding similar reports received from the sponsor as soon as possible and, in no event later than 15 calendar days after the investi gator discovers the information
Progress reports as requested by the IRB, but in any event, no less than on an annual basis, including the number of participants withdrawing from the study and the reasons for each withdrawal
Any significant protocol deviation that considerably affects the safety of the participant, or the scientific quality of the study
A final report

Additional requirements of a PI include the following:

- The investigator should be committed to a trial before the IRB issues its written approval.
- The investigator should not deviate from, or initiate, any changes to the protocol, without prior written approval from the IRB for an appropriate amendment, except when necessary to eliminate immediate hazards to the participants or when changes involve only logistical or administrative aspects of the trial (e.g., a change of the monitor telephone number).
- The investigator should promptly report the following to the IRB:
  1. Deviations from, or changes in, the protocol to eliminate immediate hazards to the trial participants.
  2. Changes that increase the risk to participants or affect significantly the conduct of the trial.
  3. All adverse drug reactions that are both serious and unexpected.
  4. New information that may adversely affect the safety of the participants or the conduct of the trial.

The investigator should be aware that the IRB has the authority to suspend or terminate approval of research that is not being conducted in accordance with IRB requirements or FDA regulations or that has been associated with unexpected serious harm to participants.

In cases in which an investigator engages in serious or continuing noncompliance, it is the responsibility of the IRB to report that activity to the sponsor and to the FDA. Noncompliance issues include, but are not limited to, unreported changes in the protocol, misuse or nonuse of the informed consent document, failure to submit protocols to the IRB in a timely manner, and avoiding or ignoring the IRB.

22.9 Research Participants

Research participants are what drive the entire clinical trial process. Without a sufficient number of volunteers, statistically significant conclusions about new drugs and devices would not be possible. Given that research volunteers have a wide variety of medical knowledge, there are mechanisms in place to ensure subjects are able to make educated decisions regarding study participation. The informed consent form alleviates the need for specialized knowledge on the part of the volunteer. In lay terms, the informed consent form describes, in detail, potential risks and benefits.

Regulations also exist to protect the confidentiality of the research participant. All information collected throughout the clinical trial remains with the study staff. For the purposes of data capture, each subject is identified by initials or study number only. In addition, the informed consent discusses who will have access to the trial documents.

Research participants are the true pioneers of medicine. Through their participation, novel therapeutic cures and treatments have been made possible. Furthermore, their participation also protects the public from approval of drugs that have a poor benefit-to-risk relationship. Thus, data obtained via research volunteers may be used to provide medical advances or to protect against insidious drugs entering the marketplace.

22.10 Industry Trade Organizations and Support Services

The pharmaceutical industry must concede to a high failure rate. To minimize failures owing to poor performance on the part of any niche service providers, the need for structured education, training, and communication is clear. The largest trade organizations devoted to the clinical trial industry include the Associates of Clinical Research Professionals (ACRP, www.arcpnet.org), the Drug Information Associates (DIA, www.diahome.org), and the Pharmaceutical Research and Manufacturers Association (PhRMA, www.phrma.org). In addition, the FDA offers many training sessions and hosts a web site for consumers as well as researchers. *CenterWatch* is the leading publication (and a leading web site, www.centerwatch.com) for clinical trial information and industry news. Publications, such as the *Good Clinical Practice Handbook*, the *Code of Federal Regulations*, and the *ICH Guidelines* outline industry-specific standards and regulations. These publications are also available at no cost online.

22.11 Summary: Ethical Dilemmas in Clinical Research

The rapid growth in the clinical research industry is met by a similar increase in knowledge, medical innovations, new technologies, and scientific breakthroughs. Yet despite these goods, many ethical dilemmas have also been raised. Such dilemmas as ethical violations, conflicts of interest, coercion, and misrepresentation date back to the 1930s. Ethical breaches include the study of untreated syphilis conducted by the United States Public Health Services from 1932 to 1972, in which patients were not told they had syphilis, were not offered effective treatment, and were not allowed to be drafted because then they would receive treatment in the military. Clearly, these human subjects were deceived, coerced, and treated unjustly. In Nuremburg, Germany, 23 doctors were charged with crimes against humanity for performing medical experiments on concentration camp inmates and others without their consent. In the 1950s' Willowbrook incident, mentally retarded children were deliberately infected with the hepatitis virus. The parents were coerced when they were told that if their child enrolled into the study, the child could occupy one of the very few beds in the hospital. In the 1960s, live cancer cells were injected into 22 senile patients at the Jewish Chronic Disease Hospital. These patients were not told and were incapable of understanding the experiment. Between 1986 and 1990, there were as many as 3000 women with high-risk pregnancies were involved in experiments by the University of South Florida and Tampa General Hospital without their consent, a case that was settled for $3.8 million.[29] Other cases in which ethical or regulatory lapses have been sited include the 1999 death of an 18-year-old man participating in a gene therapy trial at the University of Pennsylvania and the death of a healthy 24-year-old woman participating in a clinical trial at Johns Hopkins using hexamethonium, a drug not approved by the FFA, to induce asthmalike symptoms in healthy volunteers.

Within these examples are unethical acts, coercion, conflicts of interest, and misrepresentation. Conflicts of interest are not limited to the rewards and stock options of investigators; they involve the entire research endeavor. Recently, financial conflicts of interest in medical research along with widely publicized episodes of scientific misconduct have been brought to the public's attention. In some episodes, researchers have been accused of falsifying or fabricating research data on therapeutic products in which they had substantial financial interests. There have now been many steps taken to keep such actions from occurring.

In April 2000, the American Society of Gene Therapy issued new guidelines controlling conflicts of interest in research. Among the reasons for the guidelines was the discovery that a researcher of a gene therapy trial, which involved the death of an 18-year-old subject, was heavily invested in the company that was funding the research. The American Society of Gene Therapy issued a statement making it clear that financial conflicts are unacceptable. It stated, "All investigators and team members directly involved with patient selection, the informed consent process, or clinical management in a trial must not have equity, stock options, or comparable arrangements in the companies sponsoring the trials."[30]

Because of several highly publicized deaths of patients involved in experimental studies at Universities, the impetus for Congress to proceed with human research rules has become overwhelming. Recently introduced in the U.S. House of Representatives, the *Human Research Subject Protection Act of 2002*,[31] is a proposal for all persons participating in publicly and privately funded experiments to have the legal right to informed consent and to be made aware of researchers' conflicts of interest. The proposal applies the "Common Rule" to all public and private research conducted at hospitals and academic medical centers or by CROs. It would require researchers to disclose their conflicts of interest to the patients, as well as the IRB overseeing the conduct of the clinical trial. Likewise, experts acting on IRB panels would have to report their financial ties with the industry to academic institutions. This bill includes the extension of federal oversight mechanisms, more uniformity in human protection standards, and harmonization between the rules and regulations that govern privately funded research and those that govern the publicly funded research. The bill also suggests voluntary accreditation for IRBs, provides resources for IRBs, and encourages improved training and education of investigators and IRB members.

CROs are increasingly conducting drug company research privately. The number of private practice-based investigators has grown almost fourfold within a 5-year span. This change is a response to the pressures of being the first to get a new drug to market. So, managing conflicts of interest outside of academic settings increasingly take place on a national level, as opposed to following guidelines set by an academic center. On the national level, efforts are being made to control this issue. Sponsors now have their investigators as well as any other individuals involved in their trials complete financial disclosures. They are also requiring statements about which IRB members abstained from voting (for a conflicting interest) on the approval letter from the IRB.

Publicity surrounding these activities is often inflamed rather than informative, showing clinical research in a poor light instead of offering the public the big picture of clinical research, the realities of the present, and the possibilities of the future. Thousands of investigators and their staffs who participate in clinical trials are dedicated to advancing science and developing new therapeutic treatments, and they use the highest professional standards. Reporting the positive side of clinical research along with the negative publicity about events such as those described, the health care consumer would have more complete information to make an educated choice about participating in clinical trials. The clinical research community and federal government have worked hard to create and enforce guidelines and standards of practice for the protection of human subjects. ACRP offers training programs and certifications for CRCs and CRAs and is initiating certification programs for investigators.

The research community is charged with effectively protecting human subjects and ensuring that research is conducted ethically. Ethical violations in research are often caused by lack of awareness rather than malice. However, without clinical research, medical innovations and scientific breakthroughs would not be possible. So, practicing ethical conduct, seeking continued training, and complying with the governing regulations will promote good, sound, ethical research, which will, in turn, benefit society.

We have the many years of research and the hard work of the researchers to thank for providing the treatments that we now have available for the many diseases that plague our society. Keeping the trust, interest, and confidence of the public will allow for continued, successful clinical research with new scientific developments and breakthroughs.

References

1. Orzechowski, W. and Walker, R., Dose of Reality: How Drug Price Controls Would Hurt Americans, Policy Paper No. 25, February 7, 2000.
2. Office of Public Affairs, Protecting consumers, promoting public health, in *An FDA Overview: Protecting Consumers, Protecting Public Health,* FDA, Rockville, MD, 2004.
3. *Prescription Drug User Fee Act of 1992*: PL102-571, 106 Stat 4991 (October 29, 1992), primarily codified in 21 USC §379 (h).
4. General Accounting Office, *Report to the Chairman, Committee of Health, Education, Labor, and Pensions, U.S. Senate – Food and Drug Administration: Effect of User Fees on Drug Approval Times, Withdrawals, and Other Agency Activities*, GAO 92-058, U.S. Government Printing Office, Washington, DC, 2002.
5. Tislow, J., Nesbitt, L., and Belcher, A., *Drug Injury. Liability, Analysis and Prevention,* Lawyers and Judges Publishing Company, Tucson, AZ, 2001, p. 70.
6. FDA, food and drugs: additional safeguards for children, *Code of Federal Regulations*, Title 21, Part 50, Subpart D, Rev. April 2003, p. 293.
7. Cauchom, D., Americans pay more, here's why, *USA Today*, November 10, 1999, p. A1.
8. Kolassa, E.M., The cost of research: generic drug firms get a free ride, *Washington Times*, December 6, 1999, p. A17.
9. ACRP, Where we are and where we are going, 2000 White Paper, Association of Clinical Research Professionals, Alexandria, VA, 2000, www.acrpnet.org/whitepaper2/html/ii._contract_research_ organizations.html
10. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, *Guidance for Industry: E6 – Good Clinical Practice; Consolidated Guidance*, Geneva, Switzerland, April 1996, Available at www.fda.gov/cder/guidance/959fnl.pdf
11. ACRP, *The CRC – CRA Relationship,* 2000 White Paper, Association of Clinical Research Professionals, Alexandria, VA, 2000, www.acrpnet.org/whitepaper2/html/viii._crc_cra_relatioship.html
12. *Kefauver–Harris Amendment of 1962*, PL 87-78, 76 Stat 780, codified in section 21 U.S.C. §502(n).
13. The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research,* Department of Health, Education, and Welfare, Washington, DC, April 18, 1979.
14. FDA, Food and drugs: Institutional Review Boards, *Code of Federal Regulations*, Title 21, Part 56, Food and Drug Administration, Rockville, MD, Rev. April 2003.
15. HHS, Federal Policy for the Protection of Human Subjects, *Federal Register*, 56(117), Department of Health and Human Services, Washington, DC, 18, 1991.
16. HHS, Public welfare: Protection of human subjects, *Code of Federal Regulations*, Title 45, Part 46 Subtitle A, Department of Health and Human Services, Washington, DC, Rev. October 2000, pp. 107–125.
17. *Trials of War Criminals before the Nuremberg Military Tribunals under Control Council Law No. 10,* Vol. 2, U.S. Government Printing Office, Washington, DC, 1949, pp. 181–182.
18. World Medical Organization, Declaration of Helsinki: Ethical Principals for Medical Research Involving Human Subjects, Adopted by the 18th WMA General Assembly, Helsinki, Finland, June 1964.

19. FDA, Statement of Investigator, Form 1592, Department of Health and Human Services, Washington, DC, Rev. January 2003, Available at www.fda.gov/opacom/morechoices/fdaforms/fda-1572.pdf
20. Penslar, R.D., *Human Research Subjects: Institutional Review Board Guidebook,* Office of Human Research Protections, Washington, DC, 1993.
21. Levine, R., *Ethics and Regulation of Clinical Research,* 2nd ed., Urban and Schwarzenberg, Baltimore, MD, p. 42.
22. FDA, Protection of human subjects: Elements of informed consent, *Code of Federal Regulations*, Title 21, Part 50, Subpart b 50.25 (a) (6), Rev. April 2004, p. 314.
23. FDA, Institutional Review Boards: IRB functions and operations, *Code of Federal Regulations*, Title 21, Part 56, Section 108, Food and Drug Administration, Rockville, MD, Rev. April 2004, pp. 325–326.
24. FDA, Protection of human subjects: Exception from general requirements, *Code of Federal Regulations*, Title 21, Part 50, Subpart b 50.23, Rev. April 2004, pp. 310–312.
25. FDA, Protection of human subjects: General requirements for informed consent, *Code of Federal Regulations*, Title 21, Part 50, Subpart b 50.20, Rev. April 2004, p. 310.
26. FDA, Protection of human subjects: Exception from informed consent requirements for emergency research, *Code of Federal Regulations*, Title 21, Part 50, Subpart b 50.24, Rev. April 2004, pp. 313–314.
27. Kornetsky, S.A., IRB: 101, Physician Responsibility in Medicine and Research (PRIM&R), July 2002.
28. FDA, Protection of Human Research Subjects; Clinical Investigations Which May Be Reviewed Through Expedited Review Procedure Set forth in FDA Regulations, *Federal Register* 46(8980), January 27, 1981.
29. Medical Research Lawsuit Settled, *Tampa Tribune*, March 11, 2000.
30. ASGT, *Policy of the American Society of Gene Therapy Financial Conflict of Interest in Clinical Research*, American Society of Gene Therapy, Milwaukee, WI, adopted April 5, 2000.
31. *Human Research Subject Protections Act of 2002*, 107th Cong., 2nd session, H.R. 4697.

23

Horizons for Cancer Chemotherapy (and Nonchemotherapy)

Daniel D. Von Hoff

CONTENTS

23.1 Introduction

23.1.1 The New Horizons

The timing of the writing and publication of this chapter is propitious because it is truly a time of a new horizon for cancer chemotherapy (and nonchemotherapy or not the usual cytotoxic chemotherapy). While there is absolutely no doubt that the more conventional cytotoxic chemotherapy is still making substantial contributions to improving survival and providing clinical benefit for patients with a variety of malignancies, there truly is a new horizon in cancer therapy.[1–10] This new horizon, of course, is the use of perhaps a more targeted approach to treatment of patients with a variety of malignancies. It has been made possible by the explosion of new discoveries using molecular biology

(and other techniques).[11–13] The explosion of molecular genetic findings about tumor cells has given us insight into new ways to attack tumor cells all the way from attacking the blood vasculature that supplies the tumor, to the cell signaling pathways, to the protein degradation pathways, plus multiple other mechanistic approaches. Of note is that when such a new mechanism of action is introduced there is usually clinical activity, even sometimes dramatic clinical activity when the new compound with the new mechanism of action is tested in a clinical setting.[14–24] Specific examples of this include the discovery and use of the bcr-abl and c-kit inhibitor imatinib mesylate (Gleevec™) for treatment of patients with chronic myelogenous leukemia and those with gastrointestinal stromal tumors;[14–16] use of the monoclonal antibody to Her2/*neu* for patients with advanced breast cancer;[17] the use of epidermal growth factor receptor interactive agents for patients with non-small cell lung cancer;[18–21] the dramatic results with the monoclonal antibody against CD20, rituximab, in patients with low-grade lymphomas;[22,23] and the very dramatic effect of adding a monoclonal antibody against vascular endothelial growth factor (VEGF) to chemotherapy for patients with advanced colorectal cancer (a nearly 5 month increase in median survival), just to name a few.[24]

Thus the horizon for new therapies for patients with cancer is not the same old conventional cytotoxic chemotherapies, although there is clearly still a need for those types of agents (see below). The new horizon is rather a more cytostatic approach with biologics and perhaps a more targeted approach (although as will be seen below many "targeted" approaches are anything but targeted).

23.1.2 The New Challenges (and Opportunities) for the Drug Developer

The new horizons for cancer therapy certainly will present new challenges for the drug developer. More specifically, while cytotoxic agents are usually expected to cause tumor shrinkage (a complete or partial response-by-response evaluation criteria in solid tumors), cytostatic agents are much more likely to cause stable disease with a very low response rate (e.g., tumor shrinkage by conventional criteria).[25] The fact that the new agents might not cause shrinkage of tumor very often presents the drug developer with a whole new challenge to try to document whether or not a new agent has clinical activity.[30] Shrinkage of a tumor has long been regarded as a surrogate for activity of a new agent and indeed even a method to obtain an accelerated approval for a new agent.[26,27] Taking tumor shrinkage away as a surrogate by which a drug may be approved presents a very considerable challenge to clinical trial design for the cytostatic agents. Various ways people have used or proposed to tackle these challenges are outlined below.

23.2 Agents Recently Approved, Their Mechanism of Action, and Their Trial Design for Approval

23.2.1 Conventional Cytotoxic Agents

Table 23.1 details all of the recent Food and Drug Administration (FDA)-approved agents. In this section we will detail some interesting aspects of the approach to their development. As can be seen in Table 23.1 there are 21 new entities. Of note is that only seven (33%) of the agents including capecitabine, liposomal doxorubicin, temozolomide, and oxaliplatin, ABI-007, liposomal cytarabine, and pemetrexed could be considered conventional cytotoxic

TABLE 23.1
Example of Agents Recently Approved by the Food and Drug Administration (Last 7 Years)

New Agent	Mechanism	Tumor Type	Method for Initial Approval
Capecitabine (Xeloda™)	Cytotoxic nucleoside	Breast (then colorectal)	Phase II. On the basis of response rate. Then randomized Phase III for survival
Denileukin diftitox (Oltax™)	Immunotoxin conjugate	Relapsed/refractory cutaneous T-cell lymphoma	Phase II
Alemtuzumab (Campath)	MoAb to CD52	Third-line B-CLL	Phase II
Pegulated liposomal doxorubicin (Doxil)	Liposome encapsulated doxorubicin	Ovarian	Phase II data
Temozolomide (Temodar™)	DNA-interactive	Refractory anaplastic astrocytoma	Phase II data
Ibritumomab tiuxetan (Zevalin)	MoAB radiolabeled	Relapsed refractory follicular NHL	Phase II response
Oxaliplatin (Eloxatin)	DNA-interactive	Colorectal	Randomized Phase III
Gemtuzumab ozogamicin (Mylotarg™)	Immunotoxin conjugate	Second-line AML 60 years or older	Phase II response
Anastrozole (Arimidex™)	Aromatase inhibitor	Adjuvant breast	Randomized Phase III
Gefitinib (Iressa)	EGFR tyrosine kinase inhibitor	Third-line tumor — non-small cell lung	Phase II
Imatinib Mesylate (Gleevec)	Tyrosine kinase inhibitor bcr-abl, c-kit	Chronic myelogenous leukemia	Hematologic and cytogenetic response rate, Phase III; GIST[a]
Cetuzimab (Erbitux)	Monoclonal antibody against EGF receptor	Refractory colorectal cancer in combination with irinotecan	Randomized Phase II trial
Bevacizumab (Avastin)	Monoclonal antibody to VEGF	Colorectal	Randomized Phase III
ABI-007 (Abraxane)	Albumin-coated nanoparticle paclitaxel tubulin inhibitor	Breast cancer	Randomized Phase II trial
Erlotinib (Tarceva)	EGFR tyrosine kinase inhibitor	Refractory non-small cell lung cancer	Randomized Phase III
Bortezomib (Velcade)	Proteasome inhibitor	Third-line myeloma	Phase II response rate
Tositumomab (Bexxar)	Radiolabeled MoAb to CD20	Refractory low-grade lymphoma	Phase II
Celecoxib (Celebrex™)	Cox 2 inhibitor	Decrease in polyps in patients with familial polyposis	Phase III
Liposomal cytarabine (Depocyt)	Special formulated nucleoside	Lymphomatous meningitis	Phase II cytologic response
Pemetrexed (Alimta™)	Multitargeted antifole	Mesothelioma; non-small cell lung	Phase III
5-Azacytidine (Vidaza™)	Hypomethylating agent	Myelodysplastic syndrome	Phase III

[a] GIST — gastrointestinal stromal tumors.

agents. Thus, as can be seen in Table 23.1 the new horizon of compounds with new more cytostatic mechanisms of action is already upon us (with 67% of the recently approved agents being noncytotoxic agents).

The mechanism by which the cytotoxic agents were approved is of interest. Many were approved on the basis of response rate as a surrogate for survival (e.g., via the accelerated approval path).[26,27] Of note is the development of the cytotoxic agents, capecitabine received initial accelerated approval on the basis of the surrogate endpoint of response rate in patients with refractory breast cancer. This was later followed up with a randomized Phase III trial of docetaxel with or without capecitabine in patients with refractory breast cancer with this trial demonstrating an improvement in survival for the combination of docetaxel + capecitabine vs. docetaxel alone.[8]

Liposomal doxorubicin also received initial approval on the basis of response rate in patients with refractory cancer followed up with a randomized trial of the agent vs. topotecan.[27,28] The alkylating agent temozolomide was approved in a Phase II trial for treatment of patients with anaplastic astrocytoma.[27] Of note is that, although not yet filed for full approval, temozolomide has been recently added on to radiation therapy (e.g., a major trial of radiation therapy with or without temozolomide) for patients with glioblastoma multiforme. The result of this trial showed a major improvement in survival for patients treated with the radiation plus temozolomide vs. those treated with radiation therapy alone (median survival 15 vs. 12 months, $p < 0.0001$; and 1 year survival 26 vs. 8%, $p < 0.0001$).[29]

The DNA-interactive agent oxaliplatin had a slightly different trial design leading to accelerated approval and that was used as a comparator (oxaliplatin and fluorouracil + leucovorin vs. fluorouracil + leucovorin).[31] The approval was on the basis of response rate but additional studies clearly demonstrated that oxaliplatin improved survival and indeed in an adjuvant setting led to a significant increase in disease-free survival (78.2% disease-free survival at 3 years vs. 72.99%, $p = 0.002$).[1–3]

A much more conventional approach was required for the approval of pemetrexed for the treatment of patients with advanced mesothelioma. In a trial of cisplatinum plus pemetrexed vs. cisplatinum alone, there was a significant improvement in patient survival (a median of 13 vs. 10 months).[4] In addition to an improvement in survival, the trial also had patient-oriented endpoints such as amount of dyspnea and an improvement in vital capacity (e.g., clinical benefit). There is no doubt that future clinical trials should build upon these important patient-oriented endpoints, particularly trials in which response (tumor shrinkage) may not be as commonly seen as the response with cytotoxic agents.[32,33]

Carcinomatous meningitis is a serious problem for patients. Oncologists do not have many therapeutic options. A liposomal depot form of cytarabine called Depocyt™ (AstraZeneca, Wilmington, DE, U.S.A.) was approved on the basis of a randomized trial of the agent vs. standard unencapsulated cytarabine.[27] The unique endpoint for the trial was a greater incidence of cytology response (a surrogate endpoint) in the liposomal encapsulated arm of that study. The FDA has required a follow-up trial of Depocyt vs. intrathecal Cytarabine for patients with lymphomatous meningitis vs. intrathecal methotrexate for patients with carcinomatous meningitis.

Finally, one of the most recent and most interesting approvals of a new formulation of paclitaxel was the approval of the agent ABI-007 (Abraxane™) (Abraxis Oncology, Schaumburg, IL, U.S.A.) (paclitaxel protein-bound particles for injectable suspension [albumin-bound]).[34–36] It is indicated for treatment of breast cancer after failure of combination chemotherapy for metastatic disease or relapse within 6 months of adjuvant chemotherapy. Prior therapy should have included an anthracycline unless clinically contraindicated. The endpoint for that pivotal trial of Abraxane vs. paclitaxel was a

protocol-specific target lesion response rate on the basis of data from the first six cycles with an independent radiological review of the response assessment by the investigators. The response rate for Abraxane was 21.5% (95% CI 16.9 to 26.7%) compared with 11.1% (95% CI 6.94 to 15.8%) for paclitaxel. No premedications were required for Abraxane in comparison to the usual premedications required for paclitaxel. Hypersensitivity reactions were seen in 4% of the patients on Abraxane vs. 12% of the patients receiving paclitaxel. The approval of this agent was so special because the *strategy* used to obtain approval was to use the FDA regulations regarding new formulations that Abraxane could be considered as a different formulation of an already approved drug. The activity would then have to be documented to be at least the same. In the case of Abraxane it was better.

All of the above cytotoxic agents have utilized a variety of approaches for approval trials and clearly demonstrate that cytotoxic agents have a great deal to offer patients.

23.2.2 New Agents with More Unique Cytostatic Mechanisms of Action and Their Unique Clinical Trial Designs

Table 23.1 has 14 other agents which have been relatively recently approved by the FDA and have some unique card, for most of them are noncytotoxic mechanisms of action. As can be seen in the table there are:

- Three monoclonal antibodies
- Three tyrosine kinase inhibitors
- Two immunotoxin conjugates
- Two radiolabeled monoclonal antibodies
- One proteasome inhibitor
- One Cox-2 inhibitor (approved for prevention of polyp formation)
- One hypomethylating agent

Since Table 23.1 covers the last 7 years it is of significance to note that how much the mechanisms of action have shifted to biologic agents and signal transduction (tyrosine kinase) inhibitors. Also creeping into the mix are the radiolabeled monoclonal antibodies as well as the immunotoxin conjugates. From an historical perspective it is quite remarkable how quickly the agents with new mechanisms of action have appeared and how they have had a significant enough clinical impact to be approved.

As noted above, the major challenge with agents with new mechanisms of action, particularly cytostatic agents, is that the usual surrogate marker endpoint of response cannot be easily used.[37] However, this is certainly not the case for all of the cytostatic agents when particular trial designs are used to put that usual endpoint of response in a proper clinical perspective (i.e., make sure that the response is meaningful for an individual patient even if the response rate is low). Of course, cytostatic agents are more likely to be less toxic than cytotoxic agents, which clearly allow tolerance of a lower response rate (e.g., more patients exposed to the agent neither receive benefit from the agent nor suffer from side effects).

23.2.2.1 Monoclonal Antibodies

The three monoclonal antibodies approved, noted in Table 23.1, in the last 7 years for treatment of patients with cancer include Alemtuzumab (Campath™) (Genzyme Corporation, Cambridge, MA, U.S.A.), Cetuzimab (Erbitux™) (ImClone Corporation, Branchburg, NJ, U.S.A.), and bevacizumab (Avastin™) (Genentech, South San Francisco, CA, U.S.A.).

Campath was given accelerated approval on the basis of response rate. The pivotal trial was a study in 93 patients who had been previously treated with an alkylating agent and whose disease had progressed on treatment with fludarabine.[27] The FDA requested a follow-up study for full approval of Campath by randomizing patients between Campath and chlorambucil (Leukeran) as front-line therapy. This trial is ongoing. Thus, Campath has had a conventional development approach.

The monoclonal antibody to VEGF Bevacizumab (Avastin) also has had a fairly conventional development program. Figure 23.1 details the design of the pivotal trial for approval of Avastin.[24] The results were quite dramatic. This trial design was reported on the heels of another trial design in which patients with renal cell carcinoma (Von Hippel–Lindau deleted — so their tumor expressed VEGF) were randomized to high-dose Avastin vs. low-dose Avastin vs. placebo. The endpoint for the study was time to tumor progression (TTP). Of importance is that Avastin had so few side effects that the trial could be blinded and placebo-controlled. As noted in Figure 23.2, the data and safety monitoring board stopped the trial because the TTP in the Avastin-treated patients was twice that noted for the patients receiving placebo.[38] This extremely well-designed study was so important because it clearly demonstrated that Avastin not only had biologic activity but it also demonstrated that a strategy for enrollment of patients whose tumors have a specific genetic abnormality allows smart trial design with a high probability of success.

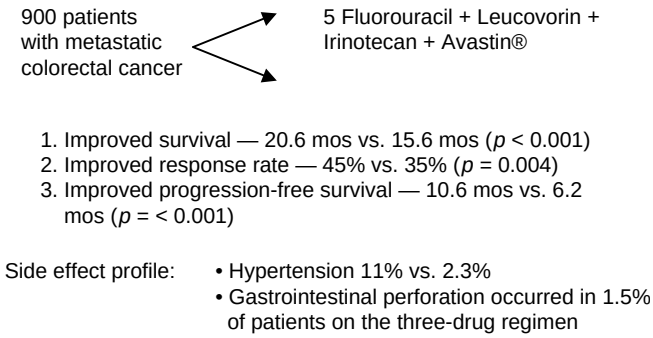


FIGURE 23.1
Trial design and result for approval of Avastin. (Reproduced from Hurwitz, H. et al., *N. Engl. J. Med.*, 350, 2335–2342, 2004.)

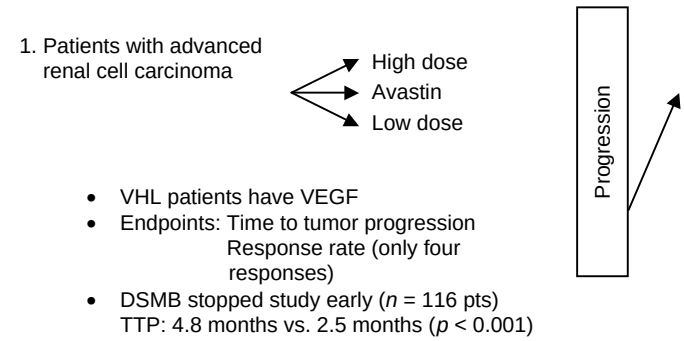


FIGURE 23.2
The clinical trial with Avastin against patients' renal cell carcinoma.

This trial demonstrates excellence in design because it enrolled patients whose tumor had expression of the VEGF (e.g., the patient's tumor had the upregulated target).[27]

Another creative trial design was performed for the pivotal trial for approval of the monoclonal antibody cetuzimab (Erbitux). The clinical trial was based on the preclinical evidence. Erbitux was synergistic in animal model systems.

As can be seen in Figure 23.3, patients who had clearly progressed on Irinotecan were randomized (2:1) to receive Cetuximab + Irinotecan (at the same dose and schedule they had just progressed on) vs. Cetuximab alone.[39] The response rate with the combination nearly doubled. The trial design clearly demonstrated a new way to design a trial if one truly believes that one's new agent is synergistic (for antitumor activity) with an agent the patients had already progressed on.

In summary, these three approvals with the noncytotoxic monoclonal antibodies demonstrated activity based on survival (Avastin), TTP in an enriched population (Avastin in renal cell cancer), a pure response with endpoint (Campath), and a clever response rate to a combination vs. a reference point response rate with that new agent alone (Cetuximab).

23.2.2.2 Tyrosine Kinase Inhibitors

The agents approved include Gefitinib (Iressa™) (AstraZeneca), Imatinib mesylate (Gleevec) (Novartis Pharmaceuticals, East Hanover, NJ, U.S.A.), and Erlotinib (Tarceva™) (Genentech).

Gleevec was the first of the tyrosine kinase inhibitors approved via an accelerated approval for patients with CML (accelerated, chronic, and blast crisis phases). This approval was based on three Phase II trials.[14,27,40,41]

One must remember that the patients were enriched for the target bcr-abl because a requirement for going on trial was that the patients had to have Philadelphia chromosome positive CML. Follow-up of these data eventually led to a full approval as did a randomized trial of Gleevec vs. cytarabine and interferon α in the first-line treatment of patients with CML. Clearly, Gleevec awakened the world to targeted therapy and gave us *a whole new horizon* for hope for our patients. Of note is that the hope was not only for patients with a hematologic malignancy. Later work using Gleevec in patients with a solid tumor, GIST, absolutely documented that a targeted therapy would also work for patients with solid tumors.[15,16] This truly changed oncology *forever*.

Gefitinib (Iressa) did not have as smooth a path for approval as Gleevec did. There were two large randomized clinical trials of chemotherapy with or without Iressa, which demonstrated no impact on survival.[42,43] These results were severely disappointing but

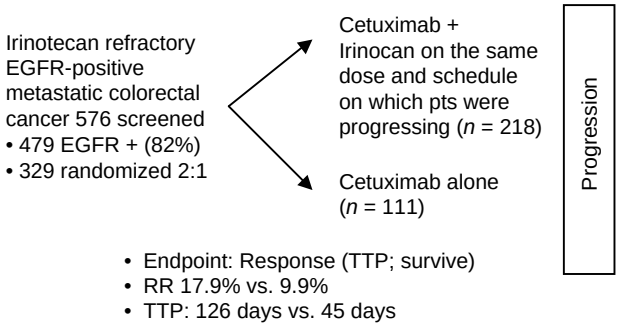


FIGURE 23.3
Design of trial for the approval of Cetuximab (Erbitux™).

were possibly due to the fact that a cytostatic agent such as Iressa could put the tumor cells in G_o and thus interfere with the carboplatin + taxol or the cisplatin + gemcitabine chemotherapy. However, because Iressa did give a real (but low) response rate in patients with non-small cell lung cancer, particularly in patients with bronchoalveolar carcinoma and in women who are nonsmokers, the agent was given accelerated approval just on the basis of Phase II data.[18,19,44,45]

More recently, there has been some exciting new information indicating that the patients whose tumors clearly and in some cases dramatically responded to Iressa have activating mutations and other mutations in the epidermal growth factor receptor gene in their tumors.[46–48] This finding has very important implications whether patients' tumors should be evaluated for these types of mutations before patients are placed on Iressa. Only additional studies will define that future for Iressa.

Erlotinib (Tarceva) took a different approach to a pivotal trial. In the study of Tarceva for patients with advanced, previously treated non-small cell lung cancer,[20] where survival was the primary endpoint, patients were randomized to receive either Tarceva or placebo. Tarceva clearly prolonged overall survival (6.7 vs. 4–7 months, $p = 0.001$). In addition, Tarceva improved progression-free survival, and improved time-to-deterioration of patients reported symptoms (cough, dyspnea, and pain). Obviously, this was a definitive trial (drug vs. placebo); however, it could be a controversial trial as it was a survival trial with no provisions to crossover to the active Tarceva. It is doubtful that many trials with this noncrossover design (with a survival endpoint) will be done in the future.

23.2.2.3 Immunotoxin Conjugates

The two immunotoxin conjugates agents that were granted accelerated approval involved some clever trial designs. Ontak™ (Ligand Pharmaceuticals), an antibody to CD25 linked to diphtheria toxin was approved on the basis of a randomized trial of *two different doses* of the agent in patients with recurrent cutaneous T-cell lymphoma, which usually expresses the CD25 component of the IL2 receptor. A follow-up trial for full approval is underway.[27]

Myleotarg™ (Wyeth Laboratories, Madison, NJ, U.S.A.), an immunotoxin conjugate of an anti-CD33 monoclonal antibody conjugated to a highly cytotoxic chalechimycin was approved for treatment of elderly patients with acute myelogenous leukemia who were not candidates for chemotherapy.[49,50] Again, a Phase III follow-up trial has been deemed a necessity.[27]

The approval of these immunotoxin conjugates demonstrates rather remarkably how this truly targeted technological approach has already shown it can work in at least two clinical situations.

23.2.2.4 Other Agents

The other agents listed in Table 23.1 include the radiolabeled monoclonal antibodies Zevalin™ (BiogenIdec, Cambridge, MA, U.S.A.) and Bexxar™ (GlaxoSmithKline, Philadelphia, PA, U.S.A.), which are being used in patients with low-grade lymphoma. Zevalin was given accelerated approval on the basis of a randomized trial of Zevalin vs. rituxan.[27] The objective response rate was higher for Zevalin and thus an accelerated approval was granted. However, a follow-up Phase III trial has been requested.[27]

The other compound of great interest with a unique mechanism of action is the proteosome inhibitor bortezumib (Velcade™) (Millennium Pharmaceuticals, Inc., Cambridge,

MA, U.S.A.). The recent approval of Velcade has been based on Phase II data in which Velcade or Velcade and desamethasone gave a response rate in patients with multiple myeloma who had received at least two prior therapies (progressing on the most recent one). The response rate was 22.7% (95% CI = 21 to 35%).[51] A follow-up study for full approval is ongoing and appears promising.[52]

23.3 Agents with New Mechanisms of Action in the Pipeline

Table 23.2 details a sampling of a number of new agents with new mechanisms of action in the pipeline. As detailed in the table there is a tremendous number of new agents with unique mechanisms of action.

Because many of these agents are based on some recently discussed targets or activating mutations or deletions, there is an excellent opportunity for trying strategies for regulatory approval on the basis of specific characterizations of patients' tumors or some other patient enrichment technique.[46,47,53]

23.4 On the Distant Horizon — Challenges for Clinical Trial Design

There is no doubt that more and more agents will be developed because of their ability to hit tumor cells with a particular genetic signature. For example, recent work from Dr. Scott Kern's laboratory at Johns Hopkins University has indicated that pancreatic cancer cells with functional defects in the Fanconi anemia pathway (including BRCA-2 mutations) are

TABLE 23.2
Sampling of New Anticancer Agents with New Mechanisms of Action on the Horizon

Agent	Mechanism of Action
SU11248	Tyrosine kinase inhibitor, VEGFR, PDGFR, Kit
BMS 354825	Src inhibitor, also abl and kit
BAY 43-9006	VEGFR and RAF
CCI779, Rad 001	m-TDR inhibitors
GCS 100LE	Antigalectin 3
Pertuzuamab	MoAb to epitope of Her2/*neu*
HMN 214	Polo-kinase inhibitor
BMS 599626	Pan Erb inhibitor
17-Allylamino 17-demethoxygeldanamycin	HSP90 inhibitor
Imexon	Mitochondrial-interactive agent
Pro1762	TRAIL interactive
DN101	Calcitriol vitamin D interactive
AQ4N	Prodrug converted by hypoxic conditions
PX 478	HIF1α inhibitor
SB 15992	Kinesin inhibitor
VX 680	Aurora kinase inhibitors
CRX 026	Dual mechanism — PRL1 inhibitor and kinesin inhibitor

more sensitive to mitomycin C.[54] BRCA-2 mutations have been implicated in about 7% of pancreatic cancer cases. Therefore, one major problem of testing the hypothesis was whether mitomycin C would cause regression of patients' tumors with a BRCA-2 mutation is the sheer number of patients who would have to be evaluated (e.g., 100 patients to obtain 7 on study). It will likely be the case that we will have to screen a large number of patients to detect those who have the precise genomics needed to ensure a response of their tumor. Clinical trial systems need to be put in place to accomplish those genomic characterizations in real time.

In addition to trying to individualize patient therapy there is still a great need for additional endpoints to evaluate the efficiency of new agents. For example, the endpoint of *time to tumor progression* has been proposed. A type of endpoint such as that would allow us to conduct randomized placebo controlled trials (with crossover) (see Figure 23.4).

A recent design called the randomized discontinuation design has attracted a great deal of attention because it leads to the great interest (based on activity) in the purported *raf* kinase inhibitor BAY 43-9006.[55] Figure 23.5 details this design. This is a design which should be seriously considered for evaluating new, less toxic cytostatic agents.

Functional imaging such as PET scanning and SPECT scanning is also likely to be explored for detecting early activity of these cytostatic agents just as PET scanning documented the early activity of Gleevec in patients with GI stromal tumors.[56]

23.5 Summary

There is a tremendous amount of ongoing activity in the development of new agents for treating patients with cancer. Also not covered in this chapter are the excellent studies of

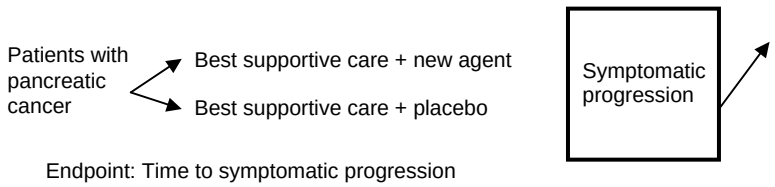


FIGURE 23.4
Schema for a time to symptomatic progression endpoint.

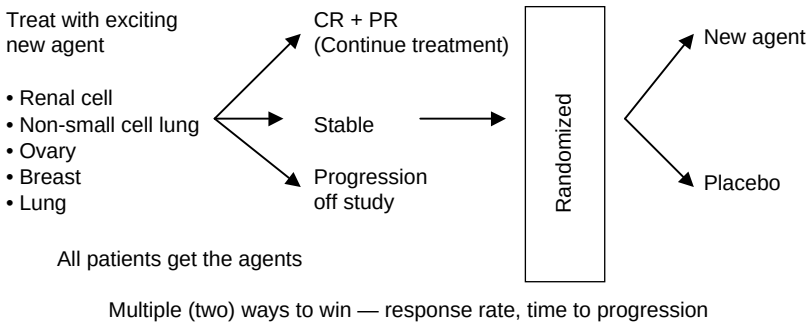


FIGURE 23.5
One form of the randomized discontinuation trial design.

new agents for the supportive care for patients with cancer. There is no doubt our burgeoning knowledge of genomics and proteomics will help make additional strides against this disease.

Acknowledgment

The author thanks Angela Wakeham for her expert editorial assistance in the preparation of this chapter.

References

1. Goldberg, R.M. et al., A randomized controlled trial of fluorouracil plus leucovorin, irinotecan, and oxaliplatin combinations in patients with previously untreated metastatic colorectal cancer, *J. Clin. Oncol*., 22, 23–30, 2004.
2. Giacchetti, S. et al., Phase III multicenter randomized trial of oxaliplatin added to chronomodulated fluorouracil-leucovorin as first-line treatment of metastatic colorectal cancer, *J. Clin. Oncol*., 18, 136–147, 2000.
3. Andre, T. et al., Oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment for colon cancer, *N. Engl. J. Med*., 350, 2343–2351, 2004.
4. Vogelzang, N.J. et al., Phase III study of pemetrexed in combination with cisplatin versus cisplatin alone in patients with malignant pleural mesothelioma, *J. Clin. Oncol*., 21, 2636–2644, 2003.
5. Hanna, N. et al., Randomized phase III trial of pemetrexed versus docetaxel in patients with non-small-cell lung cancer previously treated with chemotherapy, *J. Clin. Oncol*., 22, 1589–1597, 2004.
6. Petrylak, D.P. et al., Docetaxel and estramustine compared with mitoxantrone and prednisone for advanced refractory prostate cancer, *N. Engl. J. Med*., 351, 1513–1520, 2004.
7. Tannock, I.F. et al., Docetaxel plus prednisone or mitoxantrone plus prednisone for advanced prostate cancer, *N. Engl. J. Med*., 351, 1502–1512, 2004.
8. O'Shaughnessy, J. et al., Superior survival with capecitabine plus docetaxel combination therapy in anthracycline-pretreated patients with advanced breast cancer: phase III trial results, *J. Clin. Oncol*., 20, 2812–2823, 2002.
9. Burris, H. et al., Improvements in survival and clinical benefit with gemcitabine as first-line therapy for patients with advanced pancreas cancer: a randomized trial, *J. Clin. Oncol*., 15, 2403–2413, 1997.
10. Saltz, L.B. et al., Irinotecan plus fluorouracil and leucovorin for metastatic colorectal cancer, *N. Engl. J. Med*., 343, 905–914, 2000.
11. Hanahan, D. and Weinberg, R.A., The hallmarks of cancer, *Cell*, 100, 57–70, 2005.
12. Ferrara, N., Gerber, H.P., and LeCouter, J., The biology of VEGF and its receptors, *Nat. Med*., 9, 669–676, 2003.
13. Dancey, J. and Sausville, E.A., Issues and progress with protein kinase inhibitors for cancer treatment, *Nat. Rev. Drug Disc*., 2, 296–313, 2003.
14. Kantarjian, H. et al., Hematologic and cytogenetic responses to imatinib mesylate in chronic myelogenous leukemia, *N. Engl. J. Med*., 346, 645–652, 2002.
15. van Oosterom, A.T. et al., Safety and efficacy of imatinib (STI571) in metastatic gastrointestinal stromal tumours: a phase I study, *The Lancet*, 358, 1421–1423, 2001.
16. Demetri, G.D. et al., Efficacy and safety of imatinib mesylate in advanced gastrointestinal stromal tumors, *N. Engl. J. Med.*, 347, 472–480, 2002.

17. Slamon, D.J. et al., Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2, *N. Engl. J. Med*., 344, 783–792, 2001.
18. Kris, M.G. et al., Efficacy of gefitinib, an inhibitor of the epidermal growth factor receptor tyrosine kinase, in symptomatic patients with non-small cell lung cancer: a randomized trial, *JAMA*, 290, 2149–2158, 2003.
19. Cohen, M.H. et al., FDA drug approval summary: gefitinib (ZD1839) (Iressa) tablets, *Oncologist*, 8, 303–306, 2003.
20. Shepherd, F.A. et al., A randomized placebo-controlled trial of erlotinib in patients with advanced non-small cell lung cancer (NSCLC) following failure of 1st line or 2nd line chemotherapy. A National Cancer Institute of Canada Clinical Trials Group (NCIC CTG) trial, *Proc. Am. Soc. Clin. Oncol*., 22(Suppl.), 14S, Abstr. #7022, 2004.
21. Bonner, J.A. et al., Cetuximab prolongs survival in patients with locoregionally advanced squamous cell carcinoma of head and neck: a phase III study of high dose radiation therapy with or without cetuximab, *Proc. Am. Soc. Clin. Oncol*., 22, 489S, Abstr. #5507, 2004.
22. McLaughlin, P. et al., Rituximab chimeric anti-CD20 monoclonal antibody therapy for relapsed indolent lymphoma: half of patients respond to a four-dose treatment program, *J. Clin. Oncol*., 16, 2825–2833, 1998.
23. Coiffier, B., Rituximab in the treatment of diffuse large B-cell lymphomas, *Semin. Oncol.,* 29(Suppl. 2), 30–35, 2002.
24. Hurwitz, H. et al., Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer, *N. Engl. J. Med*., 350, 2335–2342, 2004.
25. James, K. et al., Measuring response in solid tumors: unidimensional versus bidimensional measurement, *J. Natl. Cancer Inst*., 91, 523–528, 1999.
26. O'Shaughnessy, J.A. et al., Commentary concerning demonstration of safety and efficacy of investigational anticancer agents in clinical trials, *J. Clin. Oncol*., 9, 2225–2232, 1991.
27. Dagher, R. et al., Accelerated approval of oncology products: a decade of experience, *J. Natl. Cancer Inst*., 96, 1500–1509, 2004.
28. Gordon, A.N. et al., Recurrent epithelial ovarian carcinoma: a randomized phase III study of pegylated liposomal doxorubicin versus topotecan, *J. Clin. Oncol*., 19, 3312–3322, 2001.
29. Stupp, R. et al., Concomitant and adjuvant temozolomide (TMZ) and radiotherapy (RT) for newly diagnosed glioblastoma multiforme (GBM). Conclusive results of a randomized phase III trial by the EORTC brain & RT groups and NCIC Clinical Trials Group, *Proc. Am. Soc. Clin. Oncol*., 22, Abstr. #2, 2004.
30. Cunningham, D. et al., Randomised trial of irinotecan plus supportive care versus supportive care alone after fluorouracil failure for patients with metastatic colorectal cancer, *The Lancet*, 352, 1413–1418, 1998.
31. Rothenberg, M.L. et al., Superiority of Oxaliplatin and fluorouracil-leucovorin compared with either therapy alone in patients with progressive colorectal cancer after irinotecan and fluorouracil-leucovorin: interim results of a Phase III trial, *J. Clin. Oncol*., 21, 2059–2069. 2003.
32. Booth, B., Glassman, R., and Ma, P., Oncology's trials, *Nat. Rev. Drug Disc*., 2, 609–610, 2003.
33. Simon, R. and Maitournam, A., Evaluating the efficiency of targeted designs for randomized clinical trials, *Clin. Can. Res*., 10, 6759–6763, 2004.
34. Campbell, K.J. et al., A phase I trial of ABI-007 administered weekly for three doses every 4 weeks in patients with advanced non-hematologic malignancies, *Proc. Am. Assoc. Can. Res*., 44, 1059, Abstr. R5337, 2003.
35. O'Shaughnessy, J.A. et al., Weekly nanoparticle albumin paclitaxel (Abraxane) results in long-term disease control in patients with taxane-refractory metastatic breast cancer. Breast Cancer Research and Treatment, In Lippman, M.E., Editor-in-Chief, Special Issue 27th Annual San Antonio Breast Cancer Symposium, Kluwer Academic Publishers, Dordrecht, Vol. 88, Suppl. 1, 2004, Abstr. #1070.
36. February 2005, http://www.appdrugs.com/1010705DR
37. Temple, R.J., A regulatory authority's opinion about surrogate endpoints, in *Clinical Measurement in Drug Evaluation*, Nimno, W.S. and Tucker, G.T., Eds., Wiley, New York, 1995, pp. 3–22.

38. Yang, J.C. et al., A randomized trial of bevacizumab, an anti-vascular endothelial growth factor antibody, for metastatic renal cancer, *N. Engl. J. Med*., 349, 427–434, 2003.
39. Cunningham, D. et al., Cetuximab (C225) alone or in combination with irinotecan (CPT-11) in patients with epidermal growth factor receptor (EGFR)-positive, irinotecan-refractory metastatic colorectal cancer (MCRC), *Proc. Am. Soc. Clin. Oncol*., 22, 252, Abstr. #1012, 2005.
40. Cohen, M.H., Moses, M.L., and Pazdur, R., Gleevec for the treatment of chronic myelogenous leukemia: U.S. Food and Drug Administration regulatory mechanisms, accelerated approval, and orphan drug status, *The Oncologist*, 7, 390–392, 2002.
41. Cohen, M.H. et al., Approval summary for Imatinib Mesylate capsules in the treatment of chronic myelogenous leukemia, *Clin. Can. Res*., 8, 935–942, 2002.
42. Giaccone, G. et al., Gefitinib in combination with gemcitabine and cisplatin in advanced non-small-cell lung cancer: a phase III trial-INTACT, *J. Clin. Oncol*., 22, 777–784, 2004.
43. Herbst, R.S., Giaccone, G., and Schiller, J.H., Gefitinib in combination with paclitaxel and carboplatin in advanced non-small-lung cancer: a phase III trial-INTACT2, *J. Clin. Oncol*., 22(5), 785–792, 2004.
44. Roberts, T.G. Jr. and Chabner, B.A., Beyond fast track for drug approvals, *N. Engl. J. Med*., 351(5), 501–505, 2004.
45. West, H. et al., Authors are Franklin, Gumerlock, Vance, Lau, McCoy.] Gefitinib (ZD1839) therapy for advanced bronchoalveolar lung cancer (BAC): Southwest Oncology Group (SWOG) study S0126, *Proc. Am. Soc. Clin. Oncol*., 22(14S), 620S, Abstr. #7014, 2004.
46. Lynch, T.J. et al., Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to Gefitinib, *N. Engl. J. Med*., 350(21), 2129–2139, 2004.
47. Paez, J. et al., EGFR mutations in lung cancer: correlation with clinical response to Gefitinib therapy, *Science*, 304, 1497–1500, 2004.
48. Dancey, J.E. and Freidlin, B., Targeting epidermal growth factor receptor: are we missing the mark, *The Lancet*, 362, 62–64, 2003.
49. Bross, P.F. et al., Approval Summary: Gemtuzumab Ozogamicin in relapsed acute myeloid leukemia, *Clin. Can. Res*., 7, 1490–1496, 2001.
50. De Angelo, D.J. et al., Interim analysis of a phase II study of the safety and efficacy of gemtuzumab ozogamycin (Mylotarg) given in combination with cytarabine and daunorubicin in patients <60 years old with untreated acute myeloid leukemia, *Proc. Am. Soc. Hem*., 100, 198a, Abstr. #745, 2002.
51. Richardson, P.G. et al., A Phase 2 study of Bortezomib in relapsed refractory myeloma, *N. Engl. J. Med*., 348, 2609–2617, 2003.
52. Richardson, P. et al., Bortezomib versus dexamethasone in relapsed multiple myeloma: a phase 3 randomized study, *Proc. Am. Assoc. Can. Res*., 22, 560S, Abstr. #6511, 2004.
53. Sawyers, C., Targeted cancer therapy, *Nature*, 432, 294–297, 2004.
54. Van der Heijden, M.S. et al., Functional defects in the Fanconi anemia pathway in pancreatic cancer cells, *Am. J. Pathol*., 65, 651–677, 2004.
55. Ratain, M., Preliminary antitumor activity of BAY 43-9006 in metastatic renal cell carcinoma and other advanced refractory solid tumors in a phase II randomized discontinuation trial (RDT), *Proc. Am. Soc. Clin. Oncol*., 22, 381, Abstr. #4501, 2004.
56. Joensuu, H. et al., Brief report: effect of the tyrosine kinase inhibitor STI571 in a patient with a metastatic gastrointestinal stromal tumor, *N. Engl. J. Med*., 344, 1052–1056, 2001.

24

Human Immunodeficiency Virus/Acquired Immune Deficiency Syndrome: Clinical Testing Challenges

Vincent Idemyor

CONTENTS

24.1 Introduction

The disease known as acquired immune deficiency syndrome (AIDS) was first reported in 1981, and the virus that causes it continues to create formidable challenges to the biomedical research and public health communities around the world. Globally, about 14,000 individuals are newly infected daily with one of the nine known subtypes of human immunodeficiency virus (HIV) that cause AIDS. Cumulatively, as of the year 2004, the virus has infected more than 70 million individuals, killing about 30 million and leaving another 40 million with an infection that is ultimately fatal. More than 90% of these infections have occurred in the developing world, where access to antiretroviral therapy is minimal.

24.2 Pathogenesis of HIV Infection and the Concept of Reservoir

HIV-1 binds to the cell membrane through interaction of the viral envelope glycoprotein (gp) 120 with CD4 molecules on the cell surface. HIV-1 then fuses with the cell membrane as a result of interaction of gp 120 with chemokine receptors (CCR5 or CXCR4) on the cell. CCR5

binds the macrophage-tropic, nonsyncytium inducing viral isolates while CXCR4 binds the T cell-tropic, syncytium-inducing isolates. Classified as a retrovirus, HIV's baseline genetic material is RNA. The virus is uncoated as it penetrates the cell. Viral RNA and reverse transcriptase are released into the cell cytoplasm, where HIV transcribes its RNA into viral DNA. This viral DNA is then integrated into the host cell's human DNA via an enzyme known as integrase. Once integrated, the viral DNA serves as a template for replication of viral RNA and synthesis of viral proteins and polyproteins. Viral components assemble on the inner wall of the cell membrane and bud off into new virus particles. The HIV Gag protein is critical to the budding process. During this budding process, protease cleaves long polyproteins into their component enzymes and structural proteins to yield a mature virus particle.

Persistent, high-level viral replication is now established as the motivating factor of HIV pathogenesis with about 10 billion virions (viral particles complete with RNA and envelope) in an infected person daily.[1,2]

There have been significant advances in the identification of cellular cofactors that promote viral replication after entry of the virus while several cellular proteins are steadily being identified; we have yet to graze the surface of the identity of cellular ligands for HIV-1. Tumor susceptibility gene 101 (TSG 101) and endosome-associated complexes required for transport (ESCRT) are host factors that may mediate virus release, thus begging the question whether this pathway is amenable to drug development. There are also cellular factors that inhibit viral replication. Human T cells possess a natural defense against retroviral invasion. This is a rapidly evolving area in molecular biology. Recently described is a unique antiviral pathway–cellular protein known as APOBEC 3G (apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3G, also known as CEM15).[3] APOBEC 3G is a member of the APOBEC family of editing enzymes, which is expressed in lymphocytes and macrophages. It acts as a broadly active innate intracellular antiretroviral factor, except when counteracted by the viral infectivity factor (Vif) protein of lentiviruses. Vif interacts with APOBEC 3G, and expression of Vif alone in the absence of other HIV-1 proteins is sufficient to cause depletion of APOBEC3G.[3] Some postentry restriction proteins are still being worked out. An example is the tripartite interaction motif 5α (TRIM 5α), a member of a multigene family with about 37 family members in human. While it has distinct cellular localizations, its function is not well understood. What remains unknown are factors such as which family members are active and against what viruses and the determination of their species specificity.

In the early 1998, new discoveries about the pathogenesis of HIV-1 infection soon followed, and the existence of a latent reservoir for HIV-1 in resting CD4+ T cells was established. This reservoir is a significant cause of concern because it provides a potential mechanism for the virus to persist in an infected individual despite years of effective antiretroviral therapy.

There are at least three potential reservoirs. The first phase reflects virus produced predominantly from activated CD4+ T cells. Extracellular virus particles can be trapped on specialized cells in the germinal centers of the peripheral lymphoid tissues. These cells, known as follicular dendritic cells (FDCs), are able to retain antigenic material on their surfaces for long periods of time. This reservoir declines rapidly with a half-life of about 2 weeks.[4]

Persistently infected macrophages represent a second potential reservoir for the virus. Because HIV-1 does not kill infected macrophages, these cells can continue to release virus for their normal life span. Perelson and colleagues[5] stated that in treated individuals, this phase of decline in plasma virus is due to the turnover of infected macrophages with an estimated half-life of 14 days.

The third and potentially the most significant reservoir for HIV-1 consists of resting memory CD4+ T cells carrying an integrated copy of the viral genome. These cells must

survive for long periods of time to provide the host with immunologic memory; in other words the capacity to respond rapidly to previously encountered infections. Owing to the potential of latently infected CD4+ cells to survive for months to years, this reservoir represents the major barrier to virus eradication; a notion that is being actively explored by a number of laboratories. Finzi and colleagues[6] reported the mean half-life of 43.9 months for latent reservoirs of HIV-infected cells in adults with undetectable plasma viral loads.

What we currently know about viral pathogenicity and the mechanism of viral latency represents only a small facade of what needs to be known, thereby resulting in a tremendous challenge in clinical testing.

24.3 Role of Highly Active Antiretroviral Therapy and Its Challenges in the Developing World

There are currently four classes of antiretroviral agents available for the management of HIV disease. These agents can interfere with the fusion or binding of HIV particles to host cells (known as entry inhibitors), the blocking of reverse transcriptase (known as reverse transcriptase inhibitors), and protease inhibitors (which target the viral protease, the enzyme required for the cleavage of precursor protein gag and gag-pol, thereby permitting the final assembly of the inner core of viral particles).

There is entirely a new family of compounds being developed that inhibits virus entry. This inhibition occurs during the initial stage of virus life cycle—the attachment of the virus into cell cytoplasm by fusion of the membrane. There are many different stages of viral entry, each susceptible to antagonism by specific compounds (Figure 24.1). There are compounds in clinical development that inhibit the virus attaching to CD4+ cell surface. There are compounds that bind to CCR5 that prevent the virus envelope gp attaching to the coreceptor in clinical development.

Drugs known as integrase inhibitors and maturation inhibitors are also in development. Integrase inhibitors are designed to interfere with the integration of the viral genome while the maturation inhibitors specifically block the conversion of the HIV-1 capsid precursor, CA-SP1 (p25) to mature capsid protein (p24). This blocking will result in defective core condensation and the release of noninfectious virus particles.

The current standard of care recommends the use of potent three-drug combinations,[7] which typically involves two nucleoside/nucleotide reverse transcriptase inhibitors with either a nonnucleoside reverse transcriptase inhibitor (NNRTI) or a protease inhibitor (PI). By attacking HIV infection with three drugs at one time, practitioners seek to avoid the emergence of resistance strains within the patient.

The initial combination regimens comprise two nucleoside reverse transcriptase inhibitors (NRTIs) and a PI. These combinations were well studied and are still considered the standard of care. Other combinations involving an NNRTI and two NRTIs are also used. These combinations, often referred to "PI-sparing regimens," exert a strong and durable antiretroviral effect and permit a subsequent PI-based regimen to be used if and when resistant strains of HIV emerge in the patient. Triple NRTI combinations of zidovudine/lamivudine/abacavir can also provide durable viral suppression; however, the current Department of Health and Human Services (DHHS) guidelines[7] recommend these combinations of triple-NRTI regimen as alternative therapy.

The unprecedented benefits resulting from highly active antiretroviral therapy (HAART) have been well described in the medical literature and there is agreement that symptomatic patient and those with AIDS require antiretroviral therapy. However, the

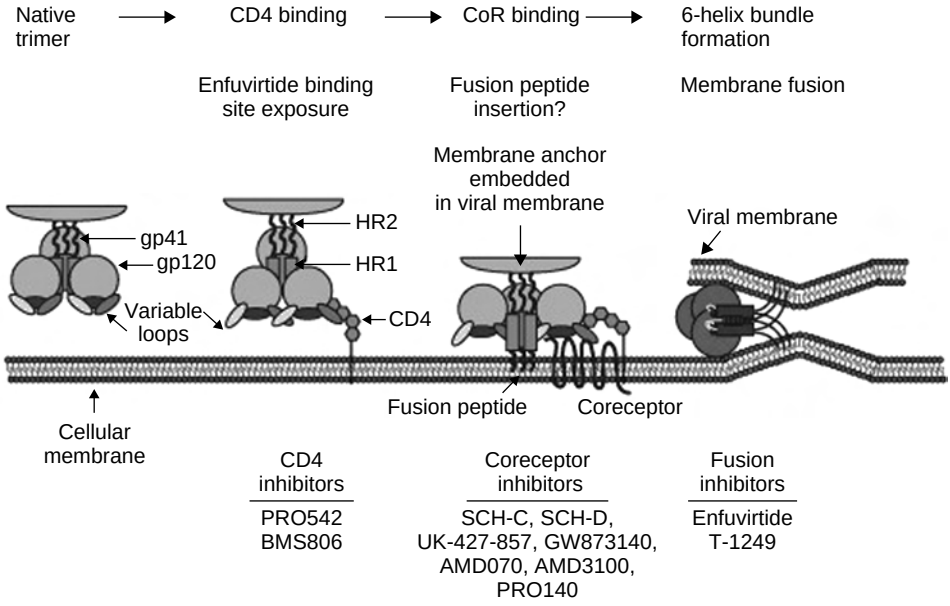


FIGURE 24.1
Components of HIV entry and targets for entry inhibition including coreceptor blockers. There are many different stages of viral entry and each of these is specifically susceptible to antagonism by specific compounds. The compounds listed above are in clinical development with the exception of enfuvirtide, which was approved by the Food and Drug Administration (FDA). (Reprinted with permission from Dr. Robert W. Doms. Available at http://clinicaloptions.com. Accessed August 25, 2005.)

timing of antiretroviral therapy in the asymptomatic patient is less clear-cut. The U.S. Department of Health and Human Services (DHHS) guidelines[7] have been revised to reflect a more conservative trend, characterized by a preference for delayed therapy and the use of CD4+ cell count as the most important criterion for initiation of therapy. These guidelines suggest that treatment be considered for individuals with CD4+ cell count less than 350 cells/mm^3 or plasma HIV-1 RNA greater than 100,000 copies/mL (by RT–PCR or bDNA assay). The point of views regarding when to initiate therapy has evolved as a result of several factors; among them is the current hypothesis that viral reservoirs do exist. This hypothesis regarding viral reservoirs has diminished the enthusiasm and rationale for early therapy. Also, evidence from multiple cohorts demonstrated that CD4+ cell count is a more significant predictor of clinical progression, mortality, and benefit from antiretroviral therapy. Patients who defer therapy (usually until CD4+ cell count has fallen to between 200 and 350 cells/mm^3) do as well as those who start therapy at an earlier stage.

The goal of initial antiretroviral therapy is suppression of viral load to below the limits of detection of plasma HIV-1 RNA less than 20 to 50 copies/mL, using an ultrasensitive viral load assay. However, it is not unusual for patients doing well on antiretroviral therapy to experience occasional "blips" of transient viremia with subsequent resuppression. The success of HAART is highly dependent upon patient adherence. Adherence is a challenge for many HIV-infected individuals owing to various reasons, including high pill burden, adverse effects, and lifestyle changes.

The availability of about 20 antiretroviral drugs, for those individuals or country programs that can afford them, in 2005 and their use in potent antiretroviral regimens has allowed for a level of control of HIV replication that was not possible before. The durability of that control, however, is threatened by the emergence of viral resistance to antiretroviral medications. Antiretroviral drug resistance occurs at a molecular level. The

nucleotide sequence of a gene determines the amino acid sequence of the protein encoded by the gene. Mutations arise during replication when changes occur to the nucleotide sequence at a given codon, resulting in the incorporation of a different amino acid that is normally found at that position in the protein.

HIV mutates rapidly averaging of 1 to 2 mutations per replication cycle. This means that in a given day, all the potential mutations that can code for single-drug resistance can easily be generated. Multiple mutations can arise in response to selective drug pressure, in which only those viral strains that possess specific mutations are able to survive in the presence of the drug.

HIV exists *in vivo* as quasi-species in which many highly related yet genetically distinct viral variants coexist. The HIV resistance testing platforms, currently available, evaluate the predominant circulating viral quasi-species. Mutant strains present in less than 20 to 30% of the viral population are less likely to be detected by testing, causing concern, and posing clinical testing challenge.

Genotypic resistance assays use DNA sequencing methods to examine the reverse transcriptase and protease regions of the HIV genome for all resistance-associated mutations. A major drawback of this testing method is that results are difficult to interpret and expert consultation is necessary.

Phenotypic resistance assays directly measure the ability of HIV-1 to replicate in a cell culture in the presence of different antiretroviral drug concentrations. This process is similar to that used to determine antibiotic resistance and is, therefore, more familiar to most clinicians. The recombinant virus, composed of a virus's reverse transcriptase and protease genes, is inserted into a standard reference strain of virus. The recombinant virus is then tested *in vitro* for the amount of drug needed to inhibit virus replication by 50%, relative to the amount of drug needed to inhibit a reference strain of virus. Phenotypic resistance testing is limited by the fact that it is conducted *in vitro* and not *in vivo*.

An emerging type of resistance testing known as virtual phenotype predicts phenotypic resistance from the genotype by using pattern recognition that is applied to large relational databases of genotypes and phenotypes.

The absence of mutational changes in these assays is not an indicator that the selected drug is successful, because success is a function of the entire regimen rather than of individual drugs, and because viral quasi-species often change rapidly and dramatically.

Research has documented that the development of antiretroviral resistance is significantly less likely to occur when a patient's viral load is maintained at an undetectable level.[8] However, reduced susceptibility in viruses from untreated patients has also been documented. Up to 18.5% of recently infected individuals have been infected with a strain of virus bearing one or more well-characterized drug resistance mutations.[9] Although HIV resistance can develop rapidly both *in vitro* and *in vivo*, the incidence of primary resistance thus appears to vary by region and with time. The development of resistance is dependent on the timing of initiation of therapy and on the stage of the disease. It is worth noting that in the context of HAART, resistance is most often the consequence of initial treatment failure.[10]

HIV isolates identified worldwide have been divided mainly into three groups: M, N, and O. Most isolates are in-group M. On the basis of the sequences of the envelope and gag genes, nine genetically distinct subtypes have been identified within the M group. Subtype B is the most prevalent form of HIV in North America and Western Europe, while subtypes A and C are the most common types in Africa and Asia. Additional information is needed on the natural susceptibility of subtypes A and C and their patterns of resistance mutations.[10] These genetically distinct subtypes pose therapeutic challenge in the design of antiretroviral therapy.

The major challenge in the developing world is the limited funding and political support coupled with insufficient capacity and understanding to effectively respond to challenges imposed by the epidemic. While the President of the United States had promised to spend $15 billion for over 5 years on AIDS in Africa and the Caribbean nations in his 2003 State of the Union speech, his budget requests fell far short of the goal. As a result, only about one million HIV-infected individuals in the developing nations are getting HAART regimens at the time of writing this chapter. The AIDS pandemic has laid bare the limits of medical science and the reality of the radically uneven distribution of economic resources. The rich world spends annually about 600 times more on defense more than Africa has on HIV/AIDS-related matters.

The management of HIV-infected individuals is now more complex than before; not only because of expanding choices of combination therapy but also the growing recognition of longer term toxicities of HAART and the clinical implications of HIV resistance. While, long-term toxicities are commonly cited as reasons for poor adherence, strategies need to be developed that will help HIV-infected individuals adhere to their medications. It is also necessary to anticipate adverse effects of these medications before they occur and develop strategies in managing them for better outcomes.

24.4 Prevention

The HIV/AIDS pandemic in the world has quickly evolved from a major health issue to a complex international emergency that undermines the social and economic fabric of nations, especially in sub-Saharan Africa, where decades of development have been reversed. In this region, nearly all the gains in life expectancy that have been made since 1950 have been wiped out. The epidemic in the Asia-Pacific region, while years behind the sub-Saharan Africa pandemic, is on a trajectory that may take it to the same point.

Unlike other microbial scourges such as malaria and tuberculosis, for which there is little that people can do to prevent infection, HIV infection in adults is almost entirely preventable by behavior modification. Extremely rare cases of contaminated blood transfusion products or contaminated hospital hypodermic needles are the exceptions. Several approaches to prevention if properly executed can be effective. As with antiretroviral therapy, combinations of approaches are needed to achieve effective prevention results.

The use of antiretroviral agents in pregnant women with HIV infection and their infants is a good and successful prevention strategy. The rate of mother-to-child transmission of HIV in the United States has been cut to negligible levels among women and infants treated with an extended regimen of some antiretroviral agents. Many developing countries lack a strategic HIV/AIDS plan. To most effectively fight HIV/AIDS, all countries must have a strategic plan in place and political leaders should be vocal proponents of science-based HIV prevention policies.

The "Treating 3 Million by 2005" plan, also known as "3 by 5," by the World Health Organization (WHO), has a target of providing antiretroviral therapy to three million infected persons meeting established medical criteria in 60 countries by the end of 2005. In December 2003 when the initiative was launched, only four countries had established antiretroviral targets in accord with 3 by 5.

Of the approximately 40 million persons currently infected with HIV worldwide, about six million are estimated to meet medical criteria for antiretroviral treatment.

WHO believes the 3 by 5 effort will strongly reinforce prevention. According to WHO, "Rolling out effective HIV/AIDS treatment is the single activity that can most effectively

energize and accelerate the uptake and impact of prevention. Under 3 by 5, this will occur as part of a comprehensive strategy linking treatment, prevention, care and full social support for people affected by HIV/AIDS. Such support is critical — both to ensure adherence to antiretroviral therapy and to reinforce prevention."[11]

24.5 Vaccines and Microbicides Development

On the global scale, HIV is recognized as several epidemics of genetically distinct types, each with characteristic geography and predominant viral strain. Owing to its dynamic nature of replication and high error rate during reverse transcription, HIV-1 has evolved into multiple subtypes, also called clades. *Intra*-subtype genetic diversity may be as high as 20%, while *inter*-subtype diversity might be as high as 35%.[12] Subtypes A, C, and D are found all over sub-Saharan Africa, whereas subtype B is dominant in the United States and most of Europe. A heterogeneous virus population normally creates difficulties in vaccine development, as seen in the development of vaccine against influenza virus infection with as low as 2% viral diversity.[12] Difficulties are greatly compounded when variability rates reach 20 to 35%.

A major scientific obstacle to the development of a safe and effective vaccine is the difficulty in establishing the precise correlates of protective immunity against HIV infection. A number of factors such as lack of an adequate animal model, lack of incentives, and genetic variability coupled with the relatively poor understanding of the immunologic mechanisms that confer protection have stymied the development of an effective HIV vaccine as of this time.

The key issues in the development of an HIV vaccine include identification of the protective immune responses to infection and the identification of the viral gene products against which the host products are directed. When HIV infects a cell, the host responds with both B-cell (humoral) and T-cell (cell-mediated) immune mechanisms. A central target of HIV infection is a group of T cells called CD4+ T lymphocytes. These CD4+ T lymphocytes are central to a successful immune response. Loss of the CD4+ T lymphocyte response is an important characteristic feature of HIV disease. Additionally, decisions about when to start treatment with a variety of drugs are most often determined by CD4+ T levels. The CD4+ T cells normally recognize foreign antigens bound to host proteins and assist B cells through the production of various cytokines in the production of antibodies. The CD4+ T cells also assist in the recruitment of another subset of T cells known as CD8+ T cells. These CD8+ T cells are commonly known as cytotoxic T lymphocytes (CTLs).

Most of the initial work with HIV vaccines was directed at developing vaccines that elicited neutralizing antibodies. Unfortunately, these neutralizing antibodies have been very narrow in the focus of their action, that is, specific almost entirely to the strain of the inoculating virus.[13] It is now apparent that developing a vaccine that elicits broad neutralizing activity is, therefore, an important quest for HIV vaccine developers.[14] However, it is very difficult to induce broadly reactive neutralizing antibodies to HIV by immunization because HIV envelope gp features loop domains with high variability. This allows the virus to evade antibody recognition. The gp 120 is also a flexible protein, assuming many novel configurations, which makes it a more difficult target for antibodies. Because HIV exhibits rapid escape from neutralizing antibodies, a number of strategies for improving neutralizing antibody responses are being researched.

Some research is now focused on the possibility of increasing the CTL responses that target virus-infected cells.[15] Even though antibodies can play an important role in

preventing the infection of susceptible cells, CD8+ CTLs remain a key element of the immune system for recognizing and lysing virus-infected cells.[16] They lyse infected cells by recruiting natural killer (NK) cells in a process known as antibody-dependent cell-mediated cytotoxicity.

HIV-specific CD8+ responses have been detected in cervicovaginal fluid from both HIV-infected women and HIV-exposed uninfected women. Because HIV usually enters the body through mucosal surfaces, most often vaginally or rectally, and replicates in lymphoid tissue, induction of immune responses at the mucosal surfaces might be an effective preventive approach. However, one major challenge to this approach is that active memory cells must be present in sufficient amounts in addition to other properties such as having enough access to infection sites.

In the past, traditional approaches to viral vaccine development have focused on using live-attenuated virus or whole-killed virus. Because of safety concerns, these methods cannot be employed in HIV vaccine development.[17] Currently, one of the big obstacles that face candidate vaccines is the inability to induce antibodies that will neutralize a broad range of primary HIV strains. Ideally, the antibodies should have broad cross-reactivity with strains of HIV isolated from infected persons in the geographic region of interest. However, controversy about both the design of assay systems to measure the neutralization of such isolates and the interpretation of the results has persisted for more than a decade.[18] There is a greater need for research on assays used to measure immune responses. At the time of writing this paper, enzyme-linked immunospot assays, in which T cells that make interferon gamma on peptide challenge, were the standard.[19] However, Hanke and colleagues[20] reported that interferon gamma may not be the best cytokine to measure because of its minimal or lack of anti-HIV effect.

Current efforts in vaccine development have predominantly utilized clade B isolates, which represent the subtype in North America and Western Europe. There is also an increased interest in the development of clade A and C vaccines for the expanding pandemic in Asia and sub-Saharan Africa. A concern in the clade-specific vaccine strategy is the potential inability to produce large amounts of vaccine specific for distinct clades. This leaves open the question of specific vs. cross-clade effectiveness. Choice of immunogen(s), adjuvant, dose, and mode of administration are also additional variables that must be addressed in candidate vaccine research.

The field of microbicide research is expanding, as several large trials among high-risk women are in progress in both the developed and the developing nations.[21] Microbicide research and development is a complex and lengthy process. Conducting vaginal microbicide trials in the developing world sometimes pose complex ethical challenges such as difficulties in obtaining informed consent, maintaining the confidentiality of the study participant's HIV status, and also difficulty in providing care to those who may be infected with HIV during the trial period.[22]

In some countries, when women are infected with HIV, they often face physical and emotional violence. As a result, they are often abandoned by their families and friends and also ostracized by their respective communities.[23] Fear of violence not only prevents women from accessing information about HIV/AIDS, it prevents them from getting tested.

It is very important that HIV/AIDS prevention programs involve both men and women so as to effectively address gender inequality. Most of the ethical precepts for experimentation on humans are designed in the developed world, where the legislative arms are in place, to some degree, to take care of the interests of the study participants. The same is not applicable in the developing world, thereby posing additional clinical challenges. At the time of writing this paper, about 40 types of microbicides are in development but none with the backing of a major pharmaceutical company.

People are dying in Malthusian numbers and most of them are now women. Globally, women comprise about half of all people living with HIV/AIDS. In sub-Saharan Africa and the Caribbean, nearly 58% of those living with HIV were women at the end of 2004. A way must be found to allow women to protect themselves, independent of male hegemony. Women must be given the liberty and control over how to protect themselves from HIV infection, and that way is microbicides. Higher levels of investment are thereby needed for effective recruitment and follow-up.

24.6 Summary and Conclusions

There are severe concerns among people of color in the wealthy nations. For example, in the United States, because of the disproportionate incidence of HIV/AIDS among people of Africa descent, special emphasis must be placed on reaching this population with effective education, treatment, and risk reduction programs despite the challenges. The current situation is so critical that not even a few people compare it to the Tuskegee medical study of untreated syphilis in African Americans, the longest nontherapeutic experiment on humans in history. Many people of sub-Saharan Africa descent are beginning to feel that the sluggish response to HIV/AIDS may be a form of genocide.

The disease known as AIDS was first reported in 1981 and the virus that causes it continues to create formidable challenges to the biomedical research and public health communities around the world. Most of the initial work on HIV vaccines was directed at developing vaccines that elicited neutralizing antibodies. These neutralizing antibodies have been narrow in the focus of their action and specific almost entirely to the strain of the inoculating virus. Additionally, controversy has been reported about both the design of assay systems to measure the neutralization of such isolates and interpretation of the results. Researchers are now looking for a "broad-spectrum" vaccine; however, the high variability of the HIV envelope gp and its rapid rate of mutation create an elusive target. Safety concerns have reduced interest in live-attenuated virus or whole-killed virus vaccines. Vaccine research considerations must include understanding the role of mucosal immunity, the importance of clades, and the continuing search for immune correlates of protection. Also, microbicides research must move forward with trials in some of the same populations that HIV vaccine trials are being considered for. Success of vaccines and microbicides will take vision, scientific breakthroughs, political will, and mobilization of far more resources than what are now made available. Given the limitations of currently available antiretroviral drugs, strategies to tackle the long-term management of HIV disease should continue to be evaluated. Research evaluating the control of HIV replication while possibly avoiding the long-term complications of the available antiretroviral agents should be encouraged.

Early detection of HIV infection is vital in controlling its spread; however, individuals in resource-limited settings feel they have no reason to find out about their HIV status if there are no treatment interventions. The availability of treatment will act as an incentive for these individuals to seek testing and counseling. There is also a strong need to create and sustain a medical and public health infrastructure in the developing world and also scale up proven prevention strategies.

Despite developments in the field of molecular biology, virology, immunology, and pharmacology, the control of HIV-1 still awaits effective vaccines and microbicides. Because significant technological advances are still required to overcome the unique obstacles posed by HIV-1, we must find ways to scale up proven prevention strategies and also provide access to HIV treatment for the infected individuals in the world.

Without an increased public health effort, the worst of the global pandemic will occur in the next decade before effective vaccines and microbicides are available.

Editor's Comment

Few would deny that acquired immune deficiency syndrome (AIDS) is one of the worst pestilences ever to beset mankind. The virus that causes this disease is incorporated into certain human cells of the immunologic system and, when activated, it destroys the ability of these cells and the body to protect itself from a wide variety of infectious parasites. In the very early 1980s, the Centers for Disease Control and Prevention (CDC) in Atlanta, GA, reported on a rare type of cancer (Kaposi's sarcoma), the frequency of which was increasing in gay men. The CDC began to warn blood banks of a possible infectious agent in the plasma processed from human donors, but actual proof was lacking. Those of us in the pharmaceutical industry involved in the production of products from human blood (e.g., antihemophilic factor, albumin, and immunoglobulin) immediately became part of a consortium of administrators and investigators at the CDC, the FDA, and the National Institutes of Health (NIH), in addition to various research and medical centers that used these products. Although a virus was suspected early on, the type of virus, if indeed one existed, was totally unknown. In conjunction with the FDA and, as this editor recalls, the CDC, the industry began experiments involving "spiking" fresh human blood with several different types of virus and processing that blood in the normal manner for the production of blood products. The ability of the surrogate viruses to survive the processing was determined. Although it was clear that the methods in place did indeed result in the death of many if not most of the added viruses, one could not be completely certain that no viable parasite existed or that the surrogate viruses chosen for the experiments were actually comparable to the ones suspected of causing AIDS.

The collaboration among the above-mentioned groups was intense and admirable. It represented a historic case in which members of the most scientific research teams in government and university laboratories collaborated totally with research scientists in the pharmaceutical industry with the objective to (1) determine whether a virus was present in human blood that might transmit this disease and (2) if so, how the products could be treated before use in human patients to eliminate the parasite. Before the end of the decade, researchers at the NIH and the Pasteur Institute in France isolated a virus they called human immunodeficiency virus (HIV), thought to be the cause of AIDS.

At the same time, and into the late 1980s, another side of industry was involved in collaborative efforts to synthesize potential drugs for the treatment of this terrible disease. As research at the NIH and various university and medical centers began to point toward HIV as the cause of the problem, there was still no clear evidence of a molecular target upon which scientists could focus to discover or develop therapeutic agents. Our experience with the treatment of viruses has always been dismal, particularly as compared to the case with bacterial infections. To this day, we cannot truly "cure" any viral infection with chemotherapy although, to be sure, some can be ameliorated. An enzyme called "reverse transcriptase" was proposed as a possible key player in the maintenance of AIDS and, as a result, programs were undertaken by industry and NIH groups to inhibit this enzyme as a valid target for chemotherapy. Thus, in a period of only approximately 5 years, researchers and experts in disease transmission had actually identified the probable cause of AIDS and related it to transmission in human blood and semen. In a period less than 10 years in duration, the first therapeutic agent for the treatment of AIDS was approved by

the FDA (zidovudine; Glaxo Wellcome, NY, U.S.A.) in 1987. At that point, the system for collection of blood was altered in the hope of reducing donations from people who have a high likelihood of being carriers of the disease and a rather broad spectrum of approaches for therapeutic agents was underway. Because of the decimation of the body's immune system by the AIDS virus, severe secondary infections by various parasites (e.g., *Pneumocystis carinii*) were identified. Again, industry and academic researchers, in conjunction with FDA, rapidly advanced a compound to treat Pneumocystis (pentamidine), which was approved by the FDA in 1989. Today, some 20 drugs have been made available for the chemotherapy of AIDS and its attendant sequelae, and extensive clinical studies have established the need to use various combinations of drugs for maximal therapeutic success. Significant improvement in life expectancy has been realized because of the availability of these drugs. To be sure, they are accompanied by toxic effects on the host, they must be used religiously and continuously and, in spite of the most careful approaches to treatment, viral resistance does develop. Nonetheless, major improvement in quality of life and longevity in AIDS patient has been achieved.

In addition to the chemotherapeutic approaches, the industry, again in collaboration with university colleagues, has been working on an effective vaccine for this disease. The history of viral disease has been one in which vaccines have been much more successful than chemotherapeutic agents, whereas, to date, the reverse is the case with AIDS.

There is no question that we still have a very long way to go to eliminate or treat AIDS in our population. Fortunately, the biotechnology field has brought us the ability to manufacture products such as "Factor VIII" by processes conducted totally outside the body, thus removing the risk of contamination with the AIDS virus. The cost of drugs is very high and prohibitive in many places in the world as was discussed by Dr Idemyor in his chapter on the subject. Those of us intimately involved in the process of new drug discovery and development and society at large can be thankful for the collaborative efforts that went into the early definition of AIDS, means of reducing the transmission into the population, and discovery of therapeutic agents. At the same time, we realize that the job is certainly not completed and tireless efforts will be required in years to come to, hopefully, bring this terrible disease to its knees.

References

1. Wei, X., Ghosh, S.K., Taylor, M.E., Johnson, V.A., Emini, E.A., Deutsch, P., Lifson, J.D., Bonhoeffer, S., Nowak, M.A., Hahn, B.H. et al.,Viral dynamics in HIV-1 infection, *Nature*, 373, 117–122, 1995.
2. Ho, D.D., Neumann, A.U., Perelson, A.S., Chen, W., Leonard, J.M., and Markowitz, M., Rapid turnover of plasma virions and CD4+ lymphocytes in HIV-1 infection, *Nature*, 373, 123–126, 1995.
3. Stopak, K., de Noronha, C., Yonemoto, W., and Greene, W.C., HIV-1 Vif blocks the antiviral of APOBEC3G by impairing both its translation and intracellular stability, *Mol. Cell.*, 3, 591–601, 2003.
4. Cavert, W., Notermans, D.W., Staskus, K., Wietgrefe, S.W., Zupancic, M., Gebhard, K., Henry, K., Zhang, Z.Q., Mills, R., McDade, H., Schuwirth, C.M., Goudsmit, J., Danner, S.A., and Haase, A.T., Kinetics of response lymphoid tissue to antiretroviral therapy of HIV-1 infection. *Science*, 276, 960–964, 1997.
5. Perelson, A.S., Essunger, P., Cao, Y., Vesanen, M., Hurley, A., Saksela, K., Markowitz, M., and HO, D.D., Decay characteristics of HIV-1-infected compartments during combination therapy, *Nature*, 387, 188–191, 1997.
6. Finzi, D., Hermankova, M., Pierson, T., Carruth, L.M., Buck, C., Chaisson, R.E., Quinn, T.C., Chadwick, K., Margolick, J., Brookmeyer, R., Gallant, J., Markowitz, M., HO, D.D.,

Richman, D.D., and Siliciano, R.F., Identification of a reservoir for HIV-1 in patients, on highly active antiretroviral therapy, *Science*, 278, 1295–1300, 1997.
7. Panel on Clinical Practices for Treatment of HIV Infection, *Guidelines for Use of Antiretroviral Agents in HIV-Infected Adults and Adolescents*, Available at http://www.AIDSinfo.nih.gov, Revision: July 15, 2005, Accessed August 23, 2005.
8. Wainberg, M.A. and Friedland, G., Public health implications of antiretroviral therapy and HIV drug resistance, *JAMA*, 279, 1977–1983, 1998.
9. Little, S.J., Holte, S., Routy, J.P., Antiretroviral resistance and response to initial therapy among recently HIV-infected subjects in North America, *Antivir. Ther.*, 6(Suppl. 1), 21, 2001.
10. Clavel, F. and Hance, A.J., HIV drug resistance, *N. Engl. J. Med.*, 350, 1023–1035, 2004.
11. World Health Organization, Treat 3 million people living with HIV/AIDS by 2005 Initiative, Available at http://www.who.int/3by5/publications/briefs/communities/en/, Accessed March 5, 2004, 2003.
12. Gaschen, B., Taylor, J., Yusim, K., Foley, B., Gao, F., Lang, D., Novitsky, V., Haynes, B., Hahn, B.H., Bhattacharya, T., and Korber, B., Diversity considerations in HIV-1 vaccine selection, *Science*, 296, 2354–2360, 2002.
13. Dolin, R., Human studies in the development of human immunodeficiency virus vaccines, *J. Infect. Dis.*, 172, 1175–1183, 1995.
14. Burton, D.R. and Montefiori, D.C., The antibody response in HIV-1 infection, *AIDS*, 11(Suppl. A), S87–S98, 1997.
15. Letvin, N.L., Barouch, D.H., and Montefiori, D.C., Prospects for vaccine protection against HIV-1 infection and AIDS, *Annu. Rev. Immunol.*, 20, 73–99, 2002.
16. Byrne, J.A. and Oldstone, M.B., Biology of cloned cytotoxic T lymphocytes specific for lymphocytic choriomeningitis virus: clearance of virus in vivo, *J. Virol.*, 51, 682–686, 1984.
17. Baba, T.W., Liska, V., Khimani, A.H., Ray, N.B., Dailey, P.J., Penninck, D., Bronson, R., Greene, M.F., McClure, H.M., Martin, L.N., and Ruprecht, R.M., Live attenuated, multiply deleted simian immunodeficiency viruses causes AIDS in infant and adult macaques, *Nat. Med.*, 5, 194–203, 1999.
18. Hanson, C.V., Measuring vaccine-induced HIV neutralization: report of a workshop, *AIDS Res. Hum. Retroviruses*, 10, 645–648, 1994.
19. Currier, J.R., Kuta, E.G., Turk, E., Earhart, L.B., Loomis-Price, L., Janetzki, S., Ferrari, G., Birx, D.L., and Cox, J.H., A panel of MHC class 1 restricted viral peptides for use as a quality control for vaccine trial ELISPOT assays, *J. Immunol. Methods*, 260, 157–172, 2002.
20. Hanke, T. and McMichael, A.J., Design and construction of an experimental HIV-1 vaccine for a year 2000 clinical trial in Kenya, *Nat. Med.*, 6, 951–955, 2000.
21. Population Council, Bill & Melinda Gates Foundation supports clinical trials of Population Council's lead candidate AIDS-fighting microbicide [press release], Available at http://biz.yahoo.com/020202/22000_1.html, Accessed April 2002, February 2, 2002.
22. Ramjee, G., Morar, N.S., Alary, M., Mukenge-Tshibaka, L., Vuylsteke, B., Ettiegne-Traore, V., Chandeying, V., Karim, S.A., Van Damme, L.; COL 1492 study group, Challenges in the conduct of vaginal microbicide effectiveness trials in the developing world, *AIDS*, 14, 2553–2557, 2000.
23. Blackstock, O., Curing stigma — The limits of antiretroviral access, *N. Engl. J. Med.*, 353, 752, 2005.

Section V

Regulatory and Legal Issues Affecting Drug Development

25

Common Technical Document: The Changing Face of the New Drug Application

Justina A. Molzon

CONTENTS

25.1 Preface

Drug approval is the goal of the long process of drug development. Once preclinical and clinical trial data have been collected, a new drug application (NDA) must be submitted to the regulatory authority for approval. Although the requirements for this submission have similarities around the world, until now, the applications have been quite different. Regulatory authorities working under the umbrella of the International Conference on Harmonisation (ICH) are hoping that development of the common technical document (CTD) will soon harmonize the application procedure and make this process simpler for applicants.

25.2 Introduction

Every drug, before receiving approval for marketing in the United States, must undergo rigorous scientific testing and scrutiny to ensure that it is safe and effective for its intended

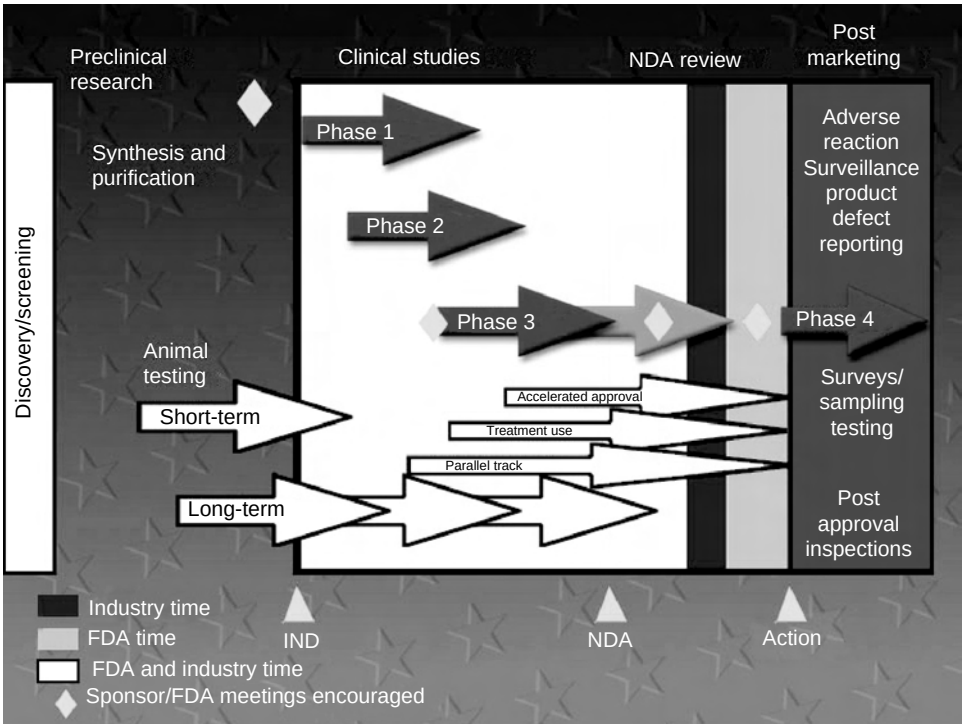


FIGURE 25.1
Investigational new drug application — submission and review process.

use. *In vivo* drug development starts with animal studies that primarily evaluate a product's pharmacology and potential toxicities. Once these studies are complete, the drug's sponsor submits an investigational new drug (IND) application for review. The IND contains the preclinical data and proposed plans for study in a human population. The information submitted in an IND is reviewed in the United States by the Food and Drug Administration (FDA), and a decision is made about whether to permit a sponsor to begin clinical trials in humans (Figure 25.1).[1]

During clinical trials, products typically undergo three phases of study: Phase 1, Phase 2, and Phase 3, which represent giving the new drug to increasing numbers of patients in an effort to characterize safety and efficacy. Once the clinical trials are completed, the results are tabulated and analyzed by the sponsor and then submitted as a NDA to the FDA. The NDA gives FDA reviewers the entire history of the drug product's development. It is FDA's job to review the data and determine if the product meets the criteria for marketing in the United States.

25.3 The Challenge

Globalization of the pharmaceutical industry has created the need to harmonize the recommendations for the development of new pharmaceuticals as well as the regulatory requirements. To address this need, in 1990, experts from the pharmaceutical industry and regulators from the United States, Europe, and Japan joined together to establish the ICH

TABLE 25.1

The Parties to the International Conference on Harmonisation (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use

Parties to the International Conference on Harmonization
European Commission — European Union (EU) The European Commission represents the 25 members of the EU. The Commission is working, through harmonization of technical requirements and procedures, to achieve a single market in pharmaceuticals, which would allow free movement of products throughout the EU.
European Agency for the Evaluation of Medicines (EMEA) The EMEA has been established by the Commission and is situated in London. Technical and scientific support for ICH activities is provided by the Committee for Proprietary Medicinal Products (CPMP) of the EMEA.
European Federation of Pharmaceutical Industries and Associations (EFPIA) The EFPIA is situated in Brussels and has, as its members, 25 national pharmaceutical industry associations and leading pharmaceutical companies. Much of the Federation's work is concerned with the activities of the European Commission and the new EMEA.
Ministry of Health, Labour and Welfare, Japan (MHLW) Affiliated with the MHLW is the National Institute of Health Sciences and Academia, which carries out research and testing on drugs, vaccines, and biologicals.
Japan Pharmaceutical Manufacturers Association (JPMA) The JPMA represents 90 member companies. Its membership includes all the major research-based pharmaceutical manufacturers in Japan.
U.S. Food and Drug Administration (FDA) The FDA has a wide range of responsibilities for drugs, biologicals, medical devices, cosmetics, and radiological products. The largest of the world's drug regulatory agencies, FDA, is responsible for the approval of all drug products used in the United States.
Pharmaceutical Research and Manufacturers of America (PhRMA) The PhRMA represents the research-based industry in the United States.
Observers
Since ICH was initiated in 1990, there have been observers associated with the process to act as a link with non-ICH countries and regions. The observers to the ICH are the World Health Organization (WHO), the European Free Trade Area (EFTA), and the Therapeutic Products Directorate of Canada.
International Federation of Pharmaceutical Manufacturers Association (IFPMA) The IFPMA is a federation of member associations representing the research-based pharmaceutical industry and other manufacturers of prescription medicines in 56 countries throughout the world. The IFPMA has been closely associated with ICH since its inception to ensure contact with the research-based industry, outside the ICH regions. Based in Geneva, Switzerland, it serves as the ICH Secretariat.

(see Table 25.1).[2] The focus of the ICH has been on the technical requirements for medicinal products containing new drugs. Because the vast majority of new drugs and medicines are developed in Western Europe, Japan, and the United States of America, when ICH was established, it was agreed that its scope would be confined to registration in those three regions.

The primary objective of the ICH is to avoid duplicative animal and human testing and to reach a common understanding about the technical requirements to support the registration process in the three ICH regions. These objectives are achieved with harmonized guidelines and the result is a more economical use of human, animal, and material resources and the elimination of unnecessary delay in the global development and availability of new medicines while maintaining safeguards on quality, safety, and efficacy and

regulatory obligations to protect public health. With the development of the CTD, the ICH hopes to accomplish many of its objectives.

There are six parties directly involved in ICH as well as observers and the International Federation of Pharmaceutical Manufacturers Association (IFPMA; see Table 25.1).[2] The six parties are the founder members of ICH, and they represent the regulatory bodies and the research-based industry in the European Union (EU), Japan, and the United States: in the EU, the European Federation of Pharmaceutical Industries and Associations (EFPIA) and the Ministry of Health, Labour, and Welfare (MHLW); in Japan, the Japan Pharmaceutical Manufacturers Association (JPMA); and in the United States, the FDA and the Pharmaceutical Researchers and Manufacturers of America (PhRMA).

25.4 Areas for Harmonization

The ICH topics are divided into four major categories: safety, quality, efficacy, and multidisciplinary and regulatory communications. Safety topics relate to *in vitro* and *in vivo* preclinical studies. Quality topics are those that relate to chemical and pharmaceutical quality assurance, and efficacy topics relate to clinical studies in human subjects. Multidisciplinary topics are ones that do not fit uniquely into any of the topic categories already mentioned and, generally, include regulatory communications. For each of the technical topics selected for harmonization, an Expert Working Group (EWG), with representatives from each of the ICH regions, reviews the differences in requirements among the three regions and develops scientific consensus to reconcile those differences. To date, ICH has approved more than 40 guidelines aimed at removing redundancy and duplication in the development and review process.

The ICH is currently working on a CTD to be used in reporting the technical requirements for a new product submission to regulatory authorities. After agreement had been reached on the various ICH topics, the next logical step was to arrange the information in a common format. This would eliminate the need for industry to reformat the information for submission to the different ICH regulatory authorities.

25.5 Components of the Common Technical Document

The CTD is organized into five modules (Figure 25.2). Module 1 is region-specific, and Modules 2 to 5 are intended to be common for all regions. More specifically, Module 1 is for administration and prescribing information and should contain documents specific to each region, for example, application forms or the proposed label for use in the region. The content and format of this module can be specified by the relevant regulatory authorities. Module 2 contains the CTD summaries and begins with a general introduction to the drug, including its pharmacologic class, mode of action, and proposed clinical use. Module 2 also provides the overall quality summary, the nonclinical overview, and the clinical overview, as well as the nonclinical written summaries, the tabulated summaries, and the clinical summary. As a foundation for the aforementioned material, Module 3 contains information on quality topics. Module 4 contains the nonclinical study reports, and Module 5 contains the clinical study reports.

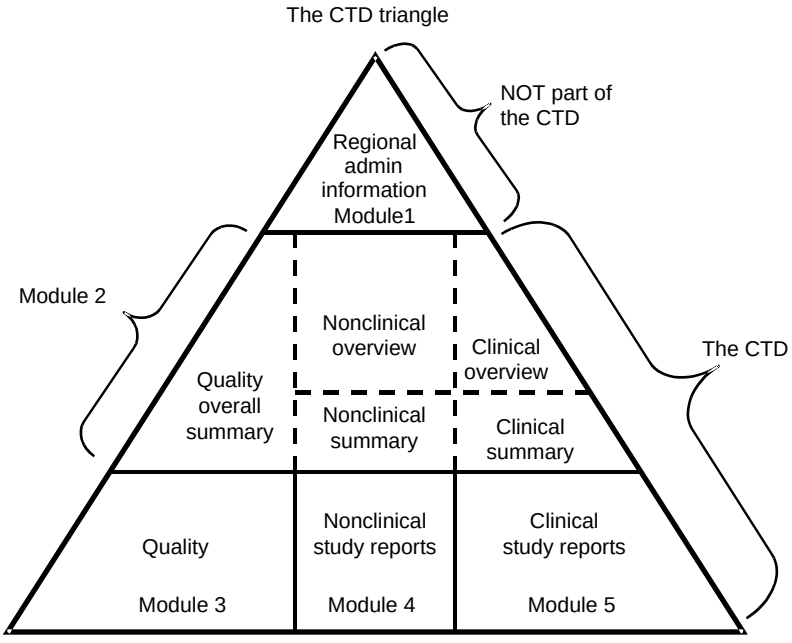


FIGURE 25.2
The five modules of the Common Technical Document.

25.6 Overcoming Inconsistencies

The CTD initiative sets out to harmonize the documentation for applications, although not to harmonize the regulatory processes for evaluating applications. Because the EU, Japan, and the United States have different regulatory systems, inconsistencies in ICH documents will remain, no matter how closely the agencies work together. However, the minor inconsistencies do not detract from the enormous achievement of the CTD. Eventually, the ambiguities and inconsistencies will be clarified and will enable the CTD to be as useful as possible. To accomplish this, the ICH regions have to work together. More importantly, regulators in the regions that have committed to the ICH format need experience with the documents to gain an understanding of which sections can be improved. While everyone is on this learning curve, receiving CTD submissions will help reviewers gain familiarity with the new format.

This is one reason why the voluntary submission phase that was initially scheduled to end in July 2002 was extended to July 2003. The extension allowed companies to become comfortable with the new format and to adapt their submission procedures accordingly. The ICH regulators have all issued guidance documents detailing what they expect to have submitted to meet the requirements of each region. On September 5, 2001, FDA's guidance on *Submitting Marketing Applications According to the ICH/CTD Format: General Considerations* was posted on the Center for Drug Evaluation and Research (CDER) website as a draft.[3] The comment period ended, officially, on November 5, 2001; by that date 12 sets of comments had been submitted. Comments are always welcome, however, especially as companies gain practical experience in assembling application data in the new format.

25.7 Are Companies Ready?

In July 2003, the CTD became the mandatory format for NDA in the EU and Japan, and the strongly recommended format of choice for NDAs was submitted to the FDA.

The ICH regulators worked with companies during the voluntary submission phase to ensure a smooth transition. One area of concern was the inclusion of the integrated summary of safety (ISS) and integrated summary of efficacy (ISE) in the CTD format. The name "summary" has caused great confusion because it implies that the data should fit within the CTD summaries in Module 2. In fact, the ISS and ISE are integrated analyses that can run to several hundred pages. They are not only a legal requirement in U.S. submissions but are also considered critical components for the safety and efficacy review. The original guidance on the ISS and ISE was issued in 1988, and the need for an updated document, taking account of the CTD, is under review.

During ICH meetings held in Washington, DC, September 9–12, 2002, CTD Implementation Working Groups developed a series of CTD Q&As (questions and answers). Intended to respond to the questions from applicants, the Q&As clarify specific areas of the CTD and are posted on the ICH website to further assist applicants in providing submission in the CTD format. The overall organization document was also updated to provide information on pagination, and the documents to be included were defined by a specific section on granularity.

Great attention was devoted in making sure that a final harmonized version of the CTD was published on the ICH website with agreed upon numbering and section headers edited for consistency and use in the electronic CTD (eCTD). National–regional versions have also been published, and the wording of the core CTD (Modules 2 to 5) may be slightly different from one region to the next because of differences in the implementation process. However, these minor changes do not affect the common understanding by ICH partners.

25.8 Achieving Perfect Harmony

Further training of reviewers and industry will be based on practical experience with submissions. Companies are therefore being encouraged to submit applications in the new format to build up this experience. The ICH Steering Committee wants to see the benefits of the CTD realized by both industry and regulators. Much work has also been done on the electronic counterpart of the CTD, the eCTD. The FDA is committed to implementation of the CTD. However implementation must be carried out as a coordinated effort. Industry and regulators must work together to make the benefits of the CTD a reality. The FDA is prepared to accept and review NDAs and biologic license applications (BLAs) in the CTD format. Industry must take the next step, and submit applications for review. The CDER received ten early submissions in CTD format as well as several submissions for new dosage forms, rolling submissions, and hybrid documents where only one of the safety, efficacy, or quality modules was submitted in CTD format. The companies making these submissions should be commended for jumping into the new process and trying it out. Furthermore, it should be noted that none of these submissions resulted in refuse-to-file actions. The various submissions were not perfect, but all were reviewable.

25.8.1 The Electronic Common Technical Document — The eCTD

Essentially, the eCTD is a transport format for facilitating electronic submissions. The eCTD serves as an interface for industry-to-agency transfer of regulatory information while at the same time, taking into consideration the facilitation of the creation, review, life cycle management, and archival of the electronic submission. The eCTD specification lists the criteria that will make an electronic submission technically valid. The eCTD represents a major advance in the submission of information to support an NDA. In the future, companies may be able to send their submissions to several regulatory authorities simultaneously with a single stroke of a computer key.

By necessity, that effort was 6 months behind the harmonized format. It is difficult to describe the specifications for electronic transmission if you do not know what you are transmitting. As a result, the eCTD Working Group had to wait for the CTD to be finalized before they could complete their work. The ICH Steering Committee adopted the final version of the eCTD at the meetings held in Washington, DC, in September 2002.

References

1. Center for Drug Evaluation and Research, *The CDER Handbook*, Food and Drug Administration, Rockville, MD, 1998, Available at www.fda.gov/cder/handbook
2. International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use, *Common Technical Document*, Geneva, 2000, Available at www.ich.org
3. Center for Drug Evaluation and Research, *Submitting Marketing Applications According to the ICH/CTD Format: General Considerations*, Food and Drug Administration, Rockville, MD, 2001, Available at http://www.fda.gov/cder/guidance/index.htm

26

Electronic Publishing

Heather L. Wallace

CONTENTS

26.1 Introduction

Electronic labeling submission of drug marketing information using the structured product labeling (SPF) format is now mandatory by the FDA since 2005. The SPF has now replaced the portable document format (PDF) as the standard acceptable submission form. The change from PDF to SPL has many advantages, one of which is to meet the new mandates in the new Medicare Prescription Drug Law. This new mandate will expedite the FDA's processing, reviewing, and archiving of the labeling content.

The new formatting process will support health information management technologies and accelerate the process of label changes. Thus, advancing the time it takes to get to the information to physician and patients. This process is only required for labeling sections or data information that has changed.

26.2 Labeling

As of June 8, 2004, the FDA made it mandatory, per the electronic labeling rule (68 FR 69009), that all labeling for marketing information be submitted electronically. Since 1999, the FDA has been able to accept electronic documents in the PDF format through Adobe Systems. The FDA is changing the type of file format it will accept for drug-label changes from PDF to the lesser known SPL format for extensible markup language (XML). It is the goal of FDA to be able to switch permanently to the SPL standard for XML format beginning in 2005. During the transition, the FDA will accept both PDF and SPL format types. The agency has determined that the PDF format does not support the new initiatives and the SPL format type has more advantages for harmonization of electronic communication.

The SPL specification is a document markup standard that specifies the structure and semantics for the regulatory requirements and content of product labeling. The SPL is derived from the HL7 clinical document architecture (CDA), which specifies the structure and semantics of "clinical documents" for the purpose of exchange.* This specification includes a detailed description of the information model for structured product labeling objects as well as the XML representation of that model.† The information model is based on the HL7 reference information model (RIM) and uses the HL7 Version 3 Data Types. (see footnote*) This version of the specification focuses on drug product labeling.

The agency noted that they are making the change to meet the electronic prescribing and health record mandates in the new Medicare prescription drug law (Public Law 108-173). To meet the new mandates, the agency has proposed to change the way it processes, reviews, and archives the content of labeling. Electronically formatted content of labeling will be used to support health information management technologies such as electronic prescribing and the electronic health record (EHR).

Electronic labeling information will improve the drug-labeling review process and speed up the approval and public dissemination of labeling changes, getting important, up-to-date information on medications to doctors and patients more quickly.

The following are SPL advantages:

- It supports the XML.
- It makes the documents both machine-readable so that they are easily parsed and processed electronically and allows the exchange of information between computer systems. It also makes them human-readable so that they can be easily retrieved and used by the people who need them.
- Only those sections or data elements of the labeling that are changed would need to be submitted rather than the complete labeling which cannot be done with PDF.
- It allows the comparison of text and specific data elements.
- It can also be used to exchange information needed for other submissions, such as drug listing, thus eliminating redundant data collection and improving efficiency.

26.3 Labeling Changes

For submission of labeling changes, only the labeling sections or data elements that have changed need to be submitted.

26.4 Conclusion

The continuous advancements made in computer technology programs to expedite information increase the access of information to its perspective audience. This enables the information to become more accessible and attainable. The submission coordinators of the FDA work with drug marketers on technical issues relating to labeling submissions.‡ The SPL and Electronic Submission Guideline documents are part of these advancements.

* The SPL specifications are available at http://www.hl7.org

† Electronic Submission Guidance documents can be found at http://www.fda.gov/cder/guidance

‡ If you have any questions or technical issues relating to a labeling submission, contact the appropriate electronic submission coordinator at esub@cder.fda.gov or esubprep@cder.fda.gov

27

The Important Role of Pharmacists in a Complex Risk-Management System: Managing the Risks from Medical Product Use by Focusing on Patient Education, Monitoring, and Adverse Event Reporting

Justina A. Molzon*

CONTENTS

27.1 Introduction

Although medical products are required to be safe, safety does not mean zero risk. A safe product is one that has reasonable risks, given the magnitude of the benefit expected and the alternatives available. All participants in the medical product development and delivery system have a role to play in maintaining this benefit–risk balance by making sure that products are developed, tested, manufactured, labeled, prescribed, dispensed, and used in a way that maximizes benefit and minimizes risk.[1]

As one of her first initiatives after being sworn in as FDA Commissioner, Dr. Jane Henney established a Task Force to evaluate the system for managing risks of FDA-approved medical products. The Task Force assessed risk-management practices within the overall healthcare-delivery system, focusing on the roles and responsibilities of each participant. The Task Force's report *Managing the Risks from Medical Product Use: Creating a Risk Management Framework* found that a systems framework for medical product risk

* The views expressed in this chapter are those of the author and do not necessarily represent the views of the U.S. Food and Drug Administration.

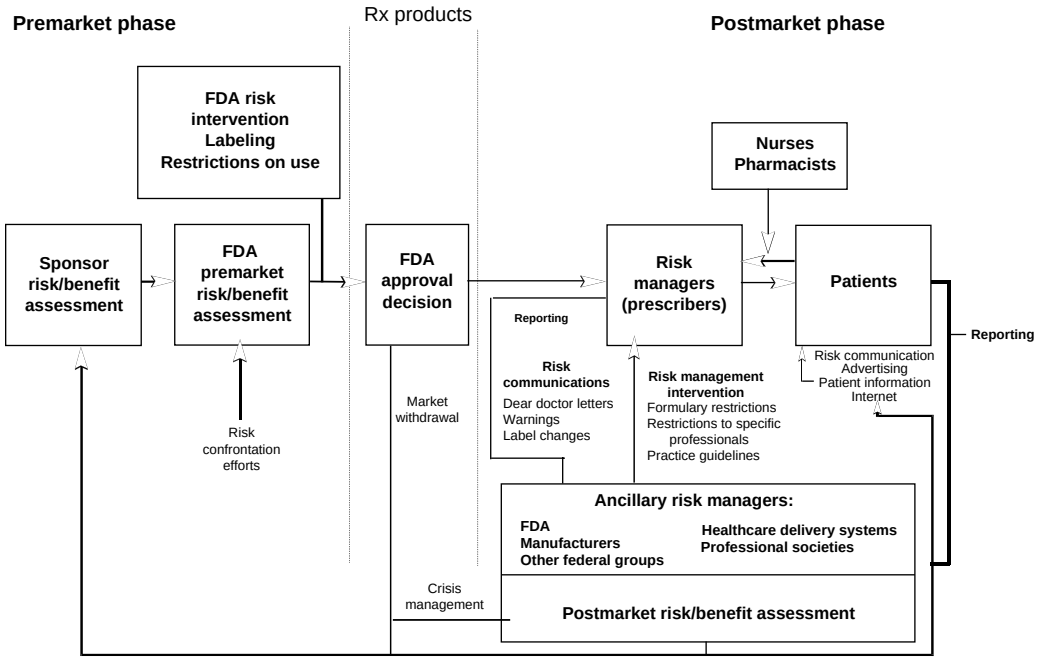


FIGURE 27.1
Complex system for managing the risks of medical products.

management would be beneficial. A systems framework should enable a better integration of the efforts of all involved parties. Such a framework should also facilitate a better understanding of both the risks involved in using medical products and the sources of those risks and enable more effective risk interventions.

As illustrated in Figure 27.1, medical products are developed and used within a complex system involving a number of key participants:

- manufacturers, who develop and test products and submit their applications for their approval to the FDA;
- the FDA, which has an extensive premarketing review and approval process and uses a series of postmarketing surveillance programs to gather data on and assess risks;
- other participants in the healthcare delivery system, including pharmacists and other healthcare practitioners; and
- patients, who rely on the ability of this complex system to provide them with needed interventions while protecting them from injury.

27.2 Pharmaceutical Care

The concept of integrating healthcare professionals' expertise for the patient's best interest is not a new one for pharmacists. Speaking at a 1989 conference focusing on evolving pharmacy practice for the 21st century, Hepler and Strand reviewed the alarming extent of drug-related morbidity and mortality in the American healthcare system. They

concluded that this problem could only be addressed by fundamentally changing the pharmacist's function, a concept they referred to as "pharmaceutical care."[2]

Defining "pharmaceutical care," as "the responsible provision of drug therapy for the purpose of achieving definite outcomes that improve a patient's quality of life," they argued that the costly social problem of "drug misadventuring" could be reduced or even eliminated by pharmacists' intervention. Rather than restricting the pharmacist's professional role to merely supplying and monitoring drug therapy, Hepler and Strand built upon the concept of clinical pharmacy to create "a process in which a pharmacist cooperates with the patient and other health professionals in designing, implementing, and monitoring a therapeutic plan that will produce specific therapeutic outcomes for the patient."

27.3 Increased Medical Errors

A recent report by the Institute of Medicine (IOM), *To Err is Human, Building a Safer Health System*,[3] claims that anywhere from 44,000 to 98,000 people die each year as a result of preventable medical errors. Many of these adverse events[†] are associated with the use of pharmaceuticals and are potentially preventable. The IOM estimates that in the United States more than 7000 deaths occur annually as a result of preventable medication errors. In addition, preventable medication errors are estimated to increase hospital costs by about $2 billion nationwide.[4]

The IOM report concludes that most of these errors are the result of systemic problems rather than poor performance by individual providers. Pharmacists play a key role in helping prevent and eliminate medication errors by providing information on the proper use of medications. Pharmacists provide the link between prescribers and patients, as shown in Figure 27.1, and can monitor for medication errors and adverse events.

27.4 Patient Education

Incorrect drug use occurs because essential information is not properly communicated – or is incompletely or incorrectly understood. Two results of not properly educating patients about their medications are mismedication and noncompliance, both of which cause unnecessary illness and health costs. Because patients usually interact more frequently with pharmacists than with their physicians, pharmacists are well positioned to monitor patients for mismedication and noncompliance. Pharmacists can ensure that a patient is taking their medication correctly and can help reduce unnecessary human suffering and increasing healthcare costs.

Patients need information to use drugs properly, not only adequate directions for use, but also information on the risks and benefits. Pharmacists are key players in the patient education process. In an article in American Pharmacy in January of 1992, former FDA Commissioner, Dr. David Kessler challenged pharmacists to renew their commitment to

[†] A number of terms are used to describe an adverse event, including adverse drug reaction (ADR), adverse experience, adverse effect, and albeit rarely, drug misadventure. In this paper, the term adverse event is used in most cases to avoid confusion.

patient education.[5] In that article he stated, "pharmacists are the 'gatekeepers at the end of a complex drug distribution system.'"[5]

As Dr. Kessler frequently pointed out, the preferred approach for patient education requires all health professionals, including pharmacists, to become more involved and more invested in the process of patient education. Patients clearly expect more from their physicians than simply writing prescriptions and more from their pharmacists than merely filling that prescription. Patients expect physicians and pharmacists to provide counseling on the proper use of their medication. Pharmacist intervention "at the gate" can prevent incorrect drug use. However, for a pharmacist to discuss with patients each of the 2.7 billion prescriptions dispensed annually is a huge task.

27.5 FDA Efforts on Patient Education

The FDA is involved in a number of initiatives to increase patient education about the proper use of medications. These initiatives should help pharmacists improve patient education and help prevent incorrect drug use.

First, FDA has been conducting continuing research into how much and what kind of information patients get about their prescriptions from pharmacies and physicians' offices. Data from the last four patient surveys that have been conducted since 1992 are on FDA's website at www.fda.gov/cder/ddmac/research.htm. FDA has also completed research on how men and women perceive benefits and risks in patient labeling as a function of how the information is communicated. FDA hopes to use these data to better communicate prescription medication information and is preparing papers for publication based on the data from this study.

Second, FDA is working with the National Association of Boards of Pharmacy (NABP) to assess the progress toward the goal of ensuring that patients receiving new prescriptions have received useful information about those prescriptions.

Third, FDA has required patient package inserts for over 50 drugs or classes to date. For some of these drugs, like Ziagen® (abacavir sulfate, GlaxoSmithKline, London, UK), Nolvadex® (tamoxifen citrate, AstraZeneca, Wilmington, DE, U.S.A.), and Halcion® (triazolam, Pharmacia & Upjohn Company, Peapack, NJ, U.S.A.), FDA believes specific patient information is essential for the drug to be marketed and used most effectively.

And, finally, FDA works with the National Council for Patient Information and Education (NCPIE) to stimulate and coordinate private sector programs to improve patient information and education.

27.6 Adverse Event Reporting

Adverse reactions to medications and the reporting of problems with drug products are also of great concern to the FDA. Pharmacists are essential participants in dealing with this concern along with other healthcare professionals.

In 1993, FDA launched MEDWatch, a program designed to promote and facilitate voluntary reporting of serious adverse events and product problems with drugs (including biologics), medical devices, and other medical products regulated by the FDA. Through MEDWatch, the FDA hoped to mobilize the entire healthcare community behind a more

rapid, effective reporting system. FDA knows that pharmacists greatly contribute to better patient care through monitoring and reporting adverse events.

The astute individual healthcare practitioner is the critical link in the postmarketing surveillance of medical products. One reporter can indeed make a difference in a newly approved drug's postmarketing phase. Only through the diligence of individual practitioners, such as pharmacists, can FDA hope to monitor the adverse events associated with newly marketed products and determine their safety for patient use.

27.7 Conclusions

Today's pharmaceuticals are not the routine antibiotics of yesteryear. FDA is now approving more sophisticated and complicated chemical entities with complex risk/benefit profiles. Some products currently on the market, such as Accutane® (isotretinoin, Hoffmann–LaRoche, Nutley, NJ, U.S.A.), might not be there without patient information accompanying the product. Patients must understand the risks and limitations of medical products so they can use them properly. By necessity, patient education will become more critical as regulatory agencies approve more drugs with complicated regimens and interaction profiles.

Patients will need to know more, about the risks — as well as the benefits — of their medications. For these reasons, a greater role for pharmacists is anticipated in improving communication with patients, providing proper patient education, monitoring the use of medications, and reporting adverse events resulting from their use.

References

1. FDA, Report to the FDA Commissioner From the Task Force on Risk Management. Managing the Risks From Medical Product Use: Creating a Risk Management Framework, Food and Drug Administration, Rockville, MD, May 1999 (http://www.fda.gov/oc/tfrm/riskmangement.pdf).
2. Hepler, C.D. and Strand, L.M., Opportunities and responsibilities in pharmaceutical care, *Am. J. Pharm. Educ.*, 53, 15S, 1989.
3. Institute of Medicine, *To Err is Human: Building A Safer Health System*, National Academy Press, Washington, DC, 1999.
4. QulC, Report of the Quality Interagency Coordination Task Force (QulC) to the President. Doing What Counts for Patient Safety: Federal Actions to Reduce Medical Errors and Their Impact, February 2000, Rockville, MD (http://www.quic.gov/report).
5. Kessler, D.A., A challenge for American pharmacists, *Am. Pharm.*, NS32, 33–36, 1992.

28

Liability, Litigation, and Lessons in New Drug Development

James T. O'Donnell

CONTENTS

28.1 Liability: The Downside of Drug Development

Therapeutic agents, such as vaccines and antibiotics, have changed the course of modern medicine and prolonged the average life span of the population by abating diseases that used to be associated with mortality. New classes of medications contribute to further treatment and disease modification. All drugs can, however, be harmful. Table 28.1 defines how the Food and Drug Administration (FDA) differentiates "adverse drug reactions" (ADRs) from "adverse drug events" (ADEs) and other terms.

In 1992, while researching ADRs as to how drug manufacturers respond to such reports, Glenna Fitzgerald, a former FDA employee, said,[1]

> The initial response of those working within an organization — particularly product champions, both technical and commercial — is that of denial. Until this moment, the whole culture within the pharmaceutical company has been positively and energetically to promote the advantage of the drug. Thus, before accepting that the drug is associated with a potentially serious disadvantage, drug champions tend, firstly, to demand proof of causality; secondly, to seek out alternative explanations for the clinical syndrome; and thirdly, to try to implicate other agents in the same drug class to diminish the impact on the specific product.

Published responses from pharmaceutical companies to ADR reports follow a similar pattern, usually containing some of the following forms of denial:

- Never admitting a causal relationship
- Never acknowledging the validity of comments in the reports
- Listing all possible alternatives however tenuous
- Never mentioning the number of cases the company has received
- Offering a long list of references, some of which seem to have little to do with the ADR
- Implying that the problem is a class effect
- Including remarks from clinicians who have been trialists or advisors to the pharmaceutical company without acknowledging the association

It is this naysay mentality that threatens the safety of medicines and exposes pharmaceutical companies to further liability.

In an editorial in the International Journal of Risk and Safety in Medicine, M.N.G. Dukes wrote that, "… even with western industry, one still, in the 1990s, runs into serious instances where risk data have been concealed in the interests of commerce."[2] Richard Merli,[3] managing editor of Pharmaceuticals Insiders, believes that increased public concern presents drug companies with the difficult choice of either "implementing expensive risk management strategies that might involve taking a blockbuster drug off the market or risking huge court awards and irreparable public relations damage." Time will tell if this is what will happen with the blockbuster Vioxx, taken off the market in 2004, with predictions of liability costs approaching $20 billion dollars. The Vioxx withdrawal and a widespread effect on the nonsteroidal anti-inflammatories (NSAIDs) market will be addressed later in this chapter.

Merli[3] states that a proactive risk management plan reduces product liability lawsuits, but it can also reduce drug sales "for reasons that turn out to be ephemeral." He warns pharmaceutical companies, however, that being defensive — such as described by Fitzgerald[1] — is extremely risky. "Ongoing product liability litigation can damage a company's bottom

TABLE 28.1
A Glossary of Drug Misadventure Terms

Terms	Glossary
Accelerated approval	A process for speeding approval of drugs that "promise significant benefit over existing therapy for serious or life-threatening illnesses." The FDA usually approves the drug on condition that the drug company will conduct phase 4 studies on the actual clinical benefit of the drug.
Adverse drug event (or adverse drug experience) (ADE)	Any untoward medical occurrence in a patient or clinical-investigation subject administered a pharmaceutical product. It is not necessary for the event to have a causal relationship with this treatment.
Adverse drug reaction (ADR)	A negative, undesirable, or harmful reaction to a particular drug. In marketed drugs, an ADR is an unintended reaction that happens at a normal dose. In clinical trials, where the drug is not yet approved for marketing, any and all unintended and noxious reactions to a drug are considered ADRs and reported accordingly.
Advisory Committee	A panel of outside experts convened periodically to advise FDA on safety and efficacy issues about drugs and other FDA-regulated products. FDA is not bound to take committee recommendations, but usually does.
Beers list or Beers criteria	Criteria for safe medication use in older adults — people over 65 years of age, the Beers list was first issued in 1991, and the criteria have been repeatedly revised and updated. Named for Dr. M.H. Beers, principal author of the original 1991 criteria.
Brand name drug	A brand name drug is a drug marketed under a proprietary, trademark-protected name.
Clinical trial	Any investigation in humans to identify any ADRs and to assess a drug's safety and efficacy.
Deception	In clinical trials, intentionally misleading or withholding information about results or reactions in the trial.
Discontinued drug	A discontinued drug is one that has been removed from the market in the United States for reasons other than safety or effectiveness.
Drug interaction	An altered reaction of the body to one drug when another is taken as well.An interaction between a drug and another substance that prevents the drug from performing as expected. This definition applies to interactions of drugs with other drugs (drug–drug interactions) as well as drugs with food (drug–food interactions) and other substances.
Efficacy	The ability of a drug to produce beneficial effects on the duration or course of a disease.
Institutional review board (IRB)	An independent group that reviews and approves material about a clinical trial to ensure that the study is safe and effective and adheres to FDA regulations.
Kaplan–Meier survival analysis	A statistical technique used to test the significance of differences between the survival curves associated with two different treatments. It is often used to analyze survival (life vs. death) data when there are censored observations (observations that are unknown because a subject has not been in the study long enough for the outcome to be observed) or to analyze the effects of different treatment procedures.[1]
Label	The FDA approved label is the official description of a drug product and includes what the drug is used for; who should take its adverse events (side effects); instructions for use in pregnancy, children, and other populations; and safety information for the patient. Labels are often found inside drug product packaging.

(*Continued*)

TABLE 28.1 (*Continued*)

Terms	Glossary
Medication Guide	A medication guide is required by FDA when a drug poses serious health concerns. It contains information for patients' understanding of how to safely use a drug product.
MedWatch program	An FDA program to monitor ADRs by analyzing voluntary health professional reports and required drug manufacturer reports.
Morbidity	An adverse effect caused by a treatment. A state produced by any departure, subjective or objective, from a state of physiological or psychological well-being.
Off-label	The unauthorized use of a drug for a purpose not approved by the FDA.
Orphan drug	An FDA designation for a therapy developed to treat a rare disease (one which afflicts less than 200,000 people in the United States). Because there are fewer financial incentives for drug companies to develop these therapies, the U.S. government offers additional incentives to the companies (tax advantages and extended marketing exclusivity) that develop these drugs.
Over-the-counter drugs (OTC)	The FDA defines OTC drugs as safe and effective for use by the general public without a doctor's prescription.
Patient package insert (PPI)	A PPI contains information for patients' understanding of how to safely use a drug product.
Phase 4 study	After a drug has been approved by the FDA, Phase 4 studies are conducted to compare the drug to a competitor, explore additional patient populations, or to further study any ADRs.
Postmarketing surveillance	The FDA's ongoing safety monitoring of marketed drugs.
Risk–benefit ratio	Risk to individual patients compared with the potential benefits.
Safety	No drug is completely safe or without potential side effects. Before a drug is approved for marketing, tests that show a drug is "safe" under the conditions on the proposed label. "Safety," therefore, is determined case-by-case and reflects the risk–benefit ratio.
Serious adverse event (SAE)	Any ADE that is fatal, life-threatening, permanently disabling, or that results in hospitalization or prolonged hospital stays.
Review	A comprehensive analysis of clinical trial data and other information that forms the basis of FDA's decision to approve an application.
Serious adverse drug event	A serious adverse event is any untoward medical occurrence that results in death, is life-threatening, requires in-patient hospitalization, prolongs a hospital stay, results in persistent or significant disability or incapacity, or causes a congenital anomaly or birth defect.
Supplement	An application to allow a company to make changes in a product that has already been approved. The FDA must approve all-important changes (in packaging or ingredients, for instance).
Side effects	Secondary or unwanted effects. Problems that occur when treatment affects healthy cells. For instance, the common side effects of chemotherapy include fatigue, nausea, vomiting, decreased blood cell counts, and hair loss. Most treatment-related side effects can be managed.
Unexpected adverse drug reaction	A reaction that is not consistent in nature or severity with study application.

Sources: From InformeDesign, University of Minnesota, www.informedesign.umn.edu; Center Watch, www.centerwatch.com/patient/glossary.html; FDA's Drug Review Glossary, www.fda.gov/fdac/special/newdrug/bengloss.html

line for years." He recommends companies to adopt a risk management plan that closely monitors ADRs and includes a communication channel for quick decisions if serious side effects begin to show up. "It's much more to a company's advantage to make a decision to voluntarily withdraw a drug from market," says John Morris, global chair of KPMG in London, "rather than be forced to withdraw the drug."[3] Poor public relations and stiff liability verdicts for not acting quickly on ADRs or for withholding drug safety information can bankrupt a company, as was seen with asbestos and now is being experienced by several ephedra manufacturers.

For instance, in September 2004, Bayer settled 2861 product liability cases for $1.09 billion for its cholesterol medicine cerivastatin (Baycol), which was linked to 100 deaths and withdrawn from market in 2001. In July 2004, the company settled 2771 cases for $1.06 billion. Bayer still has 7577 additional cases to settle[4] (see Section 28.4.4.5 for additional information). In another example, a $1 billion jury verdict was upheld against Wyeth for its fenfluramine or dexfenfluramine and phentermine (Fen–Phen) drug combination, which was linked to primary pulmonary hypertension (PPH). Wyeth has set aside $16.6 billion to cover future liability on the drug (see Section 28.4.4.2 for more on this case).[3]

28.2 History of Drug Liability and Milestones in Drug Regulation

Liability for medical and drug injuries is not new. In 2000 BC, the Babylonians decided that a physician who caused the death of a patient should lose his hands. These days, the punishment for malpractice is not quite so cruel and unusual, but medical mistakes are often career-ending errors that can cost healthcare providers millions or billions in sanctions. Table 28.2, "History and Milestones in U.S. Drug Regulation," traces the chronology of laws, regulations, and litigation intended to ensure drug safety and efficacy in the United States.

28.2.1 History of Vaccine Regulation and Litigation

As the government has an interest in preventing the spread of diseases, vaccines have followed a somewhat different historical path than that of other drugs. Most vaccines have been hailed as major modern advances. During the 1920s, several vaccines were introduced: diphtheria and tetanus toxoids, whole-cell pertussis, and bacille Calmette–Guérin (BCG) (to protect against tuberculosis) vaccines were introduced. The chorioallantoic membrane used to culture viruses allowed a yellow fever vaccine to be developed by 1935. After World War II, many vaccines still in use today emerged, including the killed and oral polio vaccines and the measles, mumps, and rubella vaccines. But vaccine regulation and litigation has a tumultuous history.

In the early 1960s, drug companies began to lobby for government indemnity for the vaccines they developed, tested, and produced.[5] Because so many people are vaccinated at one time, particularly school-age children, ADRs from a vaccine can carry considerable liability. As more diseases have become vaccine-preventable, more ADRs have been reported.[6] In 1974, impetus for indemnity increased when the courts upheld a jury verdict of $200,000 for a child who developed polio from the Sabin live-polio vaccine.[7]

After a suspected case of the 1918 Spanish flu virus (which, in a global pandemic during World War I, affected half the world's population and killed almost 25 million people in 18 months)[8] was identified in 1976, Congress passed the *National Swine Flu Immunization Program*,[9] releasing manufacturers from the liability, so that a flu vaccine

TABLE 28.2

History and Milestones in U.S. Drug Regulation

1820	Physicians establish the *U.S. Pharmacopoeia*, the first U.S. drug standards.
1848	The *Drug Importation Act* passed to stop the entry of adulterated drugs.
1862	The Bureau of Chemistry, predecessor of the FDA, is created.
1880	Chief chemist Peter Collier recommends a national food and drug law; during the next 25 years, more than 100 such bills are introduced.
1883	Chief chemist Harvey W. Wiley (called the "Crusading Chemist" and "Father of the Pure Food and Drugs Act") campaigns for a federal drug safety law.
1902	*The Biologics Control Act* passes to ensure purity and safety in vaccines.
1906	The *Food and Drugs Act* prohibits interstate commerce of adulterated drugs.
1911	In *U.S. v Johnson*, the Supreme Court rules that the *Food and Drugs Act* does not prohibit false therapeutic claims.
1912	The *Sherley Amendment* passes to prohibit false therapeutic claims in medicines.
1914	The *Harrison Narcotic Act* requires prescriptions for products exceeding the allowable limit of narcotics and mandates increased record keeping for physicians and pharmacists.
1924	The *U.S. v 95 Barrels Alleged Apple Cider Vinegar* ruling says that any statement, design, or device on a label that may mislead or deceive is prohibited.
1927	Regulatory functions of the Bureau of Chemistry become the Food, Drug, and Insecticide Administration, which was shortened to Food and Drug Administration (FDA) in an agricultural appropriations act three years later.
1933	FDA recommends revision of the *Food and Drugs Act*, launching a 5-year legislative battle.
1937	Elixir of Sulfanilamide, containing a poisonous solvent, kills 107 people including children, dramatizing the need to enact the pending food and drug law.
1938	The *Federal Food, Drug, and Cosmetic (FD&C) Act* passes requiring new drugs be shown safe before marketing, eliminating the requirement in the *Sherley Amendment* that "intent to defraud" must be proven in misbranded drugs, requiring that safe tolerances be set for unavoidable poisonous substances, authorizing factory inspections, and adding injunctions to seizures and prosecution as remedies for violations.
1940	FDA is transferred from the Department of Agriculture to the Federal Security Agency, with Walter G. Campbell appointed the first Commissioner of Food and Drugs.
1941	The *Insulin Amendment* requires FDA to test and certify the purity and potency of insulin.
1943	*U.S. v Dotterweich* makes corporations and officials of corporations liable for violations without the need to prove intent or knowledge.
1944	The *Public Health Service Act* passes to regulate biological products, the control of communicable diseases, and other health issues.
1945	The *Penicillin Amendment* requires FDA to test and certify the safety and effectiveness of all penicillin products (later extended to all antibiotics).
1948	The *Miller Amendment* ensures the FD&C Act applies to all goods regulated by FDA and transported from one state to another.
1949	FDA publishes the first *Guidance to Industry.*
1950	*Alberty Food Products Co. v U.S.* finds that drug labels must include the drug's purpose.
1951	The *Durham–Humphrey Amendment* defines drugs that cannot be used without a prescription from a licensed practitioner.
1952	*U.S. v Cardiff* finds that factory inspection in the 1938 FD&C Act is too vague to be enforced as criminal law. FDA consumer consultants are appointed in each field district.
1953	The *Factory Inspection Amendment* clarifies FD&C and requires FDA to provide written inspection reports.
1955	The Division of Biologics Control becomes an independent entity within the National Institutes of Health (NIH) after a polio vaccine is linked to about 260 cases of polio.
1958	FDA publishes the first list of "Substances Generally Recognized as Safe" (GRAS), containing nearly 200 substances.
1962	The sleeping pill thalidomide causes birth defects in thousands of babies in Western Europe. News reports about FDA medical officer Dr. Frances Kelsey keeping the drug from U.S. markets arouse public support for stronger drug regulation. The *Kefauver–Harris Drug Amendments* pass to ensure drug efficacy and safety; manufacturers are required to prove the effectiveness of their drugs before marketing them.

(*Continued*)

TABLE 28.2 (*Continued*)

1966	FDA has the *effectiveness* of 4000 drugs (approved on safety only) evaluated. The *Fair Packaging and Labeling Act* requires all labels to be honest and informative.
1968	FDA becomes part of the Public Health Service (PHS). The *Drug Efficacy Study Implementation (DESI)* is FDA's response to the investigation on the effectiveness of drugs.
1970	*Upjohn v Finch* upholds enforcement of the 1962 drug effectiveness amendments by ruling that commercial success alone does not constitute substantial evidence of drug safety and efficacy. FDA requires the first patient package insert.
1972	Over-the-counter (OTC) drug review begins to ensure the safety, effectiveness, and appropriate labeling of drugs sold without prescription. Regulation of biologics (serums, vaccines, and blood products) is transferred from NIH to FDA.
1973	The Supreme Court upholds the 1962 drug effectiveness law and endorses FDA action to control classes of products by regulation rather than litigation.
1976	The *Medical Device Amendments* pass to ensure safety and effectiveness of medical devices including diagnostic products, requiring some quality control, premarket approval, and performance standards on some products.
1982	Tamper-resistant packaging required after the deaths from cyanide placed in Tylenol, and the *Federal Anti-Tampering Act* makes it a crime to tamper with packaged consumer products.
1983	The *Orphan Drug Act* passes, enabling FDA to promote research and marketing of drugs needed to treat rare diseases.
1984	*Fines Enhancement Laws* of 1984 and 1987 increase penalties for all federal crimes, with double fines for corporations. The *Drug Price Competition and Patent Term Restoration Act* expedites the availability of less-costly generic drugs by permitting FDA to approve generics without repeating the safety and effectiveness research.
1985	FDA approves the AIDS test on blood to protect patients from infected donors.
1986	The *Childhood Vaccine Act* requires vaccine makers to provide patient information, gives FDA authority to recall biologics, and authorizes civil penalties.
1987	Regulations on experimental drugs for patients with serious diseases and no alternative therapies.
1988	The *Food and Drug Administration Act* puts FDA within the Department of Health and Human Services (HHS) and spells out research, enforcement, and education requirements. The *Prescription Drug Marketing Act* prevents the resale of diverted drugs, requires drug wholesalers be licensed, restricts reimportation, and bans the sale, trade, or purchase of drug samples or counterfeit drug coupons.
1989	FDA recalls all OTC dietary supplements containing L-tryptophan because of a clear link to an outbreak of eosinophilia-myalgia syndrome (EMS). By 1990, more than 1,500 cases of EMS – 38 deaths are confirmed, with an estimated 3,000 to 10,000 unreported cases.
1990	The *Safe Medical Devices Act* requires reporting of medical devices that probably caused the death, serious illness, or injury of a patient. Postmarket surveillance on permanently implanted devices required with methods for tracing and locating patients depending on such devices. FDA is authorized to recall device product.
1991	Regulations are published to accelerate the review of drugs for life-threatening diseases.
1992	The *Generic Drug Enforcement Act* imposes debarment and other penalties for illegal acts involving abbreviated drug applications. The *Prescription Drug User Fee Act* (PDUFA) requires manufacturers of drug and biologics to pay fees for product applications and supplements. The *Mammography Quality Standards Act* requires facilities to be accredited and federally certified by 1994, with annual inspections.
1997	The *Food and Drug Administration Modernization Act* reauthorizes PDUFA and mandates accelerated reviews and regulates drug advertising of unapproved uses.
1998	The first phase begins of consolidating FDA labs from 19 facilities to nine (by 2014).

Note: From the beginning of civilization, people have been concerned about the quality and safety of their food and medicines. In 1202, King John of England proclaimed the first English food law, the Assize of Bread, which prohibited adulteration of bread with such ingredients as ground peas or beans. Food regulation in the United States started in colonial times. Federal control of the drug supply began with the inspection of imported drugs in 1848. The chronology describes some of the milestones in the history of food and drug regulations in the United States.

Source: From Milestones in U.S. Food and Drug Law History, www.fda.gov/opacom/backgrounders/miles.html

could be produced. Forty million people in the United States were vaccinated against this flu in less than 3 months.[10] Adverse drug reactions were reported shortly after the massive vaccination program began, and it was found that those who had received the vaccine had a tenfold increase of Guillain Barre Syndrome (GBS). The swine flu program is a milestone in vaccine litigation history as it was a precursor to federal involvement and to no-fault vaccine compensation programs.[5]

In response to continued concerns about vaccine safety, the *National Childhood Vaccine Injury* (NCVI) *Act of 1986* established a no-fault compensation process for people injured by them.[11] The NCVI also mandated that the Institute of Medicine (IOM) reviews scientific evidence of vaccine-related ADRs in children. In 1996, the Department of Health and Human Services (HHS) made changes to the NCVI, which lessened its usefulness.[5] In response to the problems this generated, the *Vaccine Injured Children's Compensation Act of 2001* was introduced in Congress.[12] This bill, however, also has its problems, and in April 2001, the bill was referred to the House subcommittee on health, where it still remains.[13]

Vaccine liability issues were also covered in Section 304 of the *Homeland Security Act* (HSA) *of 2002*,[14] as amended in April 2003,[15] in which Congress enacted liability protection for manufacturers of smallpox vaccines. Vaccine liability can be handled in four different ways: the government can substitute itself as the defendant, it can decide nobody need be liable and provide no-fault compensation, it can indemnify manufacturers after they have been sued and lost, or it can alter the normal rules of litigation.[16] In the HSA, the government substitutes itself as the defendant if the HHS Secretary declares "an actual or potential bioterrorist incident or other actual or potential public health emergency makes advisable the administration of a covered countermeasure,"[16] such as a vaccine. Secretary Tommy Thompson issued the first such declaration on January 24, 2003.

The HSA states, however, that "covered countermeasures" apply only to smallpox vaccines at this time, and although HSA protects manufacturers and others against liability, it does not directly set forth compensation procedures for vaccine recipients. Under Part C of the HSA, compensation for death benefits is capped at $262,100. If a case does end up in court, the plaintiff must "prove culpability equal to or rising above the level of negligence."[16] It seems assured that the history of vaccine regulations and litigations will continue to change for some time to come. The swine flu and other vaccine-related ADRs are discussed further in Section 28.4.4.1.

28.2.2 Noteworthy Fiscal Year 2003 FDA Activities

In the Center for Drug Evaluation and Research (CDER) report for fiscal year (FY) 2003 (October 1, 2002 to September 30, 2003),[17] the agency acknowledges that the need for drug safety in the United States must be increasingly vigilant. "As Americans are increasingly receiving the benefits of important new drugs before they are available to citizens of other countries, we must be especially vigilant in our surveillance." According to Sandy Kweder, deputy director of the office of new drugs at FDA, "We have a public that is much more concerned over drug safety. Our judgments have to reflect what risks the public is willing to accept." Kweder elaborated, "There's no rule on the percentage of adverse events considered acceptable. The risks to the public must be balanced against the benefits among the patient population."[3]

The FDA's postmarketing surveillance of drug safety in FY 2003 includes a new database on prescribed drug use (with patient identities removed). The agency uses this database of marketed drugs to "make risk assessments and decisions about the most appropriate way to manage any new risk or new perspective on a previously known

risk."[17] Some of the risk management tools FDA uses to prevent confusion and associated drug injuries include the following techniques:

- Requiring new labeling
- Changing drug names
- Changing drug packaging
- Sending out "Dear Health Care Practitioner" letters
- Communicating special risk information
- Restricting distribution programs
- Terminating product marketing

The FDA has also signed a 2-year data-mining agreement with a commercial firm to develop software tools for quantitatively analyzing drug safety data. The goal is to increase the agency's awareness and understanding of trends in ADRs. Yet the regulations, litigation, and databases do not prevent serious adverse drug effects. The CDER received more than 370,880 ADRs in FY 2003, a 13% increase over the previous year.[17]

The FDA also has a drug safety email notification system. During FY 2003, 33 drug safety alerts were sent out, and 25 to 45 notices on drug labeling changes were sent out each month. In the first 4 months of 2003, FDA placed 10 drugs on their list of "Drugs with Special Safety Restrictions," which allows distribution only from specific facilities, limits prescribing authority to physicians with special training or expertise, or requires medical tests before prescriptions are written. The CDER also uses "Medication Guides," which must be distributed to patients when drugs that pose serious health concerns are dispensed. In FY 2003, CDER approved Medication Guides for one innovator product mefloquine hydrochloride (Lariam; Roche Pharmaceuticals, Nutley, NJ), on generic lindane (shampoo and lotion; Pennex Pharmaceuticals, Morton grove, IL), and on isotretinoin products (Claravis; Barr Laboratories, Pomona, NY, and Sotret; Ranbaxy Pharmaceuticals, Gurgaon, India, are the generic versions of Accutane from Roche Pharmaceuticals, which already carries a Medication Guide).

During FY 2003, CDER reviewed more than 3000 cases of medication errors, half of which, they acknowledge, were a result of error-prone labeling. During the same period, FDA recalled 254 prescription drugs (29.0% lower than in FY 2002) and 88 over-the-counter (OTC) medicines (a 6.0% increase). Drug recalls can apply to one or several batches of a drug or to the complete withdrawal of the drug (although there were no safety-based drugs withdrawn during FY 2003). FDA estimates that about 2.5% of all drugs are recalled from the market each year for safety reasons, which represents $4.5 billion in lost annual sales.[3] Between January 1, 1994 and April 30, 2004, FDA approved 303 new drug products); seven (2.3%) were subsequently recalled.[17] The primary reasons for drug recalls included the following problems:[17]

- Content uniformity failures
- Current good manufacturing practices (cGMP) deviations
- Dissolution failures
- Generic drug or new drug application discrepancies
- Label mix-ups
- Microbial contamination of nonsterile products
- Presence of a foreign substance
- pH failures
- Stability data that did not support the expiration date
- Subpotency

In other FY 2003 drug safety highlights, FDA made 1512 GMP inspections, reviewed 51 field recommendations for regulatory action, and approved 34, which included 27 warning letters, 4 injunctions, and 3 seizures. The agency also reviewed 184 foreign establishment inspection reports, resulting in one warning letter and one import alert. The CDER issued 737 drug promotion review letters (42 regulatory action letters, 185 launch campaign letters, and 510 advisory acknowledgement or closure letters). The regulatory action letters were for prescription drug promotions determined to be "false, misleading, lacking in fair balance of the risks and benefits, or promoting a product or indication before approval."[17] Warning letters are issued for more serious or repeat violations. Examples of specific types of violative promotions included promotional exhibit hall displays, oral representations, Internet sites, and journal advertisements or sales brochures. Letters for direct-to-consumer promotion violations accounted for 254 of the letters, a 26.0% increase over FY 2002.

The FDA also promulgated an OTC and prescription medicines rule, which became final in February 2004, requiring a bar code on medicines used in hospitals to ensure that health professionals gave patients "the right drugs at the appropriate dosages and at the right time."[17] Adoption of this advanced information system has, in some hospitals, reduced medication error rates by as much as 85%. The agency estimates that the rule will help prevent 500,000 adverse events and transfusion errors and save $93 billion in health costs during the next 20 years.

Another big change in 2004 was the passage of new legislation restricting drug liability class actions. During 2004, the Bush administration had gone to court as a friend to industry to prevent lawsuits by consumers who say that they have been injured by prescription drugs and medical devices.[18] This action will be discussed more extensively in Section 28.6.

That was 2003 and 2004, and of course we now know (and will read about later in the chapter) the Vioxx recall in September, 2004. A tremendous amount of renewed pressure and criticism was directed to the FDA, and the FDA has been re-examining the Agency's methods of monitoring drug safety. Janet Woodcock MD, former director of CDER, and now Deputy Commissioner for Operations, recently gave the following remarks to the IOM:

> . . . the agency's system for ensuring the safety of drugs is "pretty much broken down" and it has known for a long time it needed to improve its system. "The keystone of the current system is the prescriber, and that person is the one who decides if the benefits of a drug outweigh the risks for that patient," Dr. Woodcock said. "This system has obviously broken down to some extent, as far as the fully informed provider and the fully informed patient."
>
> She said the drug agency had long known that it needed to improve systems for learning about problems with drugs on the market. One way to do that, she said, is to take advantage of electronic health records from managed-care organizations. She also said the FDA had proposed this better system several years ago, but Congress declined to fund it. "The bottom line is that a lot of drug safety problems are actually preventable," she said, because "most adverse events are from known side effects."
>
> (*Source*: From Harris, G. Drug Safety System is Broken, a Top FDA Official Says. *New York Times*, June 9, 2005)

28.3 Classification of Adverse Drug Reactions (ADRs)

An ADR is any unintended or undesirable response obtained from the appropriate dose of a drug or diagnostic agent. These reactions have been conventionally classified into six

different categories: overdose, intolerance, unexpected side effects, secondary effects, and idiosyncratic and hypersensitivity reactions. Although drug interactions are included with other ADRs, they are also classified into seven categories:

- Those that occur outside the body
- Those that occur at the site of entry
- Those that occur at storage sites within the body
- Those that occur at the site of action
- Those that inhibit enzymes from metabolizing
- Those that stimulate enzymes
- Those that affect drug excretion

Although these classifications have been used in both reviews and textbooks, it is unnecessary and confusing to separate ADRs from drug interactions.

Drug manufacturers are required to promptly report all serious or unexpected ADRs from spontaneous sources and from any clinical or epidemiological investigation — independent of design or purpose. It also applies to cases not reported directly to a sponsor or manufacturer (for instance, those found in regulatory ADR registries or in publications). There are situations in addition to single case reports of serious ADRs that may necessitate rapid communication to regulatory authorities, such as information that might materially influence the risk–benefit assessment of a medicinal product, or would be sufficient to consider changes in administering the product or in the overall conduct of a clinical investigation**.** Examples include clinically important increases in the rate of occurrence for an "expected," serious ADR; significant hazards to the patient population — such as lack of efficacy with a drug used to treat a life-threatening disease; or major safety findings from a newly completed animal study (such as a study on carcinogenicity).

28.3.1 Type A Reactions: Augmentation of the Pharmacological Response

Type A reactions are those that are dose-related and that arise from the normal pharmacological action of a drug. They usually result from exaggerated but otherwise normal pharmacological actions of a drug given in the usual therapeutic doses. An example of this type of reaction would be the postural and exercise hypotension (low blood pressure) in a patient taking an adrenergic neuronal blocking agent such as guanethidine (Ismelin, Ciba Specialty Chemicals), used to treat severe high blood pressure, or drowsiness from taking phenobarbital (nonproprietary), a barbiturate used as a sedative and also as an anticonvulsant. Type A reactions are largely predictable on the basis of the known pharmacological properties of a drug. Although the incidence of morbidity in the population is often high, the mortality rate is usually low.

28.3.2 Type B Reactions: Bizarre (Idiosyncratic) Responses

Type B reactions are those that represent an abnormal or novel response to a drug. They are unusual effects that would not be expected from known pharmacological actions of a drug when given at the accepted dose to a patient whose body handles the drug in a normal manner. Example of Type B reactions includes malignant hyperthermia caused by anesthesia. Immunological (allergic hypersensitivity) reactions are Type B reactions (although these would constitute medication errors if the allergy is known). Anaphylaxis — a classic Type B hypersensitivity reaction — is one of the most serious and potentially life-threatening ADRs. This IgE-mediated reaction generally occurs within 20 min of exposure to an

antigen, usually after an injection, but it can occur with any route of administration. The symptoms of anaphylaxis are produced by a variety of chemical mediators, most notably histamine.[19]Although the incidence of anaphylaxis is low, the mortality rate from these reactions — which are unpredictable and are not discovered in conventional toxicological screening — is high. The Physicians' Insurance Association of America (PIAA) medication-error study[20] identifies drug allergies as a problem that results in frequent suits against physicians.

28.4 Adverse Drug Reactions: A Major Cause of Litigation

The reported incidence and frequency of ADRs vary depending on the source of the report and on the methods used to describe the event. One reason for the discrepancy is that the cause of an ADR can be difficult to ascertain. ADRs listed in hospital admission reports have ranged from less than 1% up to 28%. Most studies report an incidence rate of 10 to 20%, reflecting the different methods used to detect and report ADRs.

28.4.1 Extent of the Problem: Incidence of Adverse Drug Reactions

In reviewing ADRs, Jick[21] determined that the incidence of ADRs ranged between 1 in 10,000 (0.01%) to 1 in 200 (0.5%). ADRs with an incidence of 1 in 10,000 would be difficult to identify at the clinical trial stage, where data are usually available on only about 1000 patients. Reactions with an intermediate frequency (>1 in 10,000 but less than 1 in 200) might be identified in postmarketing surveillance studies, whereas ADRs of low frequency (<1 in 10,000) might be verified only in cohort or case-control studies. Many important ADRs are pharmacologically unpredictable with an incidence of 1 in 10,000 or less, entailing a follow-up of a cohort of drug users. This has important implications for the sample size needed in postmarketing surveillance studies. Sample-size limitations mean that serious but infrequent reactions will not be picked up unless cohort samples are of 100,000 or more.

The mortality from ADRs has been estimated within the hospitalized patient population to be in the range of 0.01 to 0.3%. Estimates of deaths annually from ADRs have ranged from 2,000 to 140,000, and annual hospitalizations due to ADRs range from 160,000 to 1.5 million.[22] It is estimated that up to 30% of hospitalized patients experience an ADR.[23]

28.4.2 Who Is Most at Risk?

The PIAA conducted a medication-error study in 1993[20] and another in 1999[24] to analyze high-frequency and severe malpractice claims. The data collected included loss description and causation information, expense and indemnity payments, and the demographics of policyholders, claimants, and institutions. The PIAA companies insured almost 87,000 physicians in the United States, ranging from the smallest to the largest physician-owned malpractice insurers.

The PIAA found in both 1993 and 1999 that prescriptions were the second-most frequent and second-most expensive item in the claims reported. As of June 30, 1992, there were 6646 claims involving medicine prescriptions. Payments were made in 2195 of these claims, resulting in a total indemnity payment of $218.9 million, an average indemnity payment of $99,721, and a median indemnity payment of $35,000. In the 1999

report, 9,801 of the 145,287 (6.7%) closed claims were for prescription medicines, and $438 million of the $7.3 billion (6.0%) paid in indemnities were for prescription medicine. The authors of the study concluded that medication injury claims are a significant source of loss for malpractice carriers and the physicians they insure.

We can extrapolate from the PIAA data to determine who is most at risk of experiencing A DRs. More female than male patients were involved in the claims studied by PIAA[20] (ratio: 1.5 females to 1 male). Two thirds of all claims involved patients between 18 and 59 years. Although the number of claims in the 6 to 12 and 13 to 17 age brackets was predictably low, these brackets had the highest average indemnities, reflecting the serious consequences of medication errors in young patients.

The elderly represent the group taking the largest amount of drugs (one third of all medications prescribed each year)[25] and suffering the greatest number of ADRs. In a recent study, Curtis et al. found that 21% (162,370 of 765,423 subjects) of elderly Americans were being prescribed drugs that could possibly harm them (defined as drugs on the Beers list, see Table 28.1), 15% received two or more of these medicines, and 4% received three or more of the risky drugs.[25]

Life expectancy — the probability of living a longer life — has increased from about 24 years for those in the Roman Empire to 30 years at the beginning of the 19th century and 50 years at the beginning of the 20th century. Now, the worldwide life expectancy is 63 years, and the average life expectancy in the United States is 77.43 years (74.63 for males and 80.36 for females), which makes it 48th in the world (Andorra has the longest at 83.5 years and Zambia has the shortest at 37.2 years). With greater life expectancy, both the number of elderly people (currently about 365 million in the United States) and the percentage of the population that is elderly (about 12.4%) increases. This means that the number of people 75 years of age and older will rise from about 10 million (4.4% of the population) in 1980 to close to 30 million (9.1% of the total population) in 2050. Whereas the population older than 85 years was about 3 million (1% of the population) in 1980, the number is anticipated to grow to about 18 million (5.2% of the population) by 2050. Those older than 85 years represent the most rapidly growing segment of the population. With the elderly experiencing the most ADRs, these statistics mean that the rate of ADRs will increase if improvements in drug safety do not increase to match it. Routine monitoring for drug safety should include patients in various subsets (the elderly, children, and patients with specific disease states, such as kidney or liver disease) where there may be a particular liability to ADRs.

The PIAA studies[20,24] also emphasize that medication errors can cause significant injuries. A substantial percentage (42.4%) of the claims involved significant permanent injuries, with 21.1% in 1993 and 27.0% in 1999 of all claims resulting in death. Close analysis of the death claims shows that medication errors were either the direct cause or a major contributing factor in the deaths in 84.3% of the 1993 claims. The average payment for a death claim ($188,555 in 1999) was 14.0% higher than for claims not involving death. In a 1998 study of 33 million hospital admissions, the number of serious injuries from ADRs was 2.2 million; that is, 2.1% of all in-patients experienced a serious ADR.[26] Lazarou and colleagues found that 4.7% of all admissions were due to serious ADRs, and fatal ADRs occurred in 0.19% of in-patients and 0.13% of admissions. The authors concluded that 106,000 deaths occur annually from ADRs.

28.4.3 Beyond Malpractice: Reasons for the High Incidence of Adverse Drug Reactions

When exploring the cause of an ADR, many syndromes and conditions can have multiple causes that occur in only a small percentage of the population and which often have vague

TABLE 28.3

Prescription Drugs Withdrawn from Market for Safety Reasons in the United States from 1997 to 2004

Drug	Brand Name	Manufacturer	Type of Drug	Date Approved	Date Withdrawn	Timeon Market	Primary Health Risk	Estimated U.S. Sales
Valdecoxib	Bextra	Pfizer, Inc.	Cox-2 inhibitor	11/1/2002	4/7/2005	2.5 years	Serious and potentially life-threatening skin reactions	(Unavailable)
Rofecoxib	Vioxx	Merck & Co.	Cox-2 Inhibitor	5/21/1999	9/30/2004	5.33 years	Heart attack and stroke	US$2.5 billion
Cerivastatin sodium	Baycol	Bayer	Cholesterollo-wering	6/26/1997	8/8/2001	4.17 years	Rhabdomyolsis	$554 million in 2000
Rapacuronium bromide	Raplon	Organon	Anesthetic, muscle relaxant	9/19/1999	3/30/2001	7 months	Bronchospasm	$23 million
Alosetron hydrochloride	Lotronex	Glaxo Wellcome	Gastrointestinal	2/9/2000	11/28/2000	9 months	Ischemic colitis	$50.4 million
Cisapride monohydrate	Propulsid	Janssen Pharmaceuticals	Heartburn	7/29/1993	7/14/2000	7 years	Heart rhythm abnormalities	US$2.5 billion
Troglitazone	Rezulin	Warner-Lambert	Type 2 diabetes	1/29/1997	3/21/2000	2.16 years	Liver failure	US$2.1 billion worldwide
Asternizole	Hismanal	Janssen Pharmaceuticals	Antihistamine	12/19/1988	5/18/1999	6 months	Heart rhythm abnormalities	$23 million
Grepafloxacin	Raxar	Glaxo Wellcome	Antibiotic	11/6/1997	11/1/1999	2 years	Heart rhythm abnormalities	$23.5 million
Mibefradil	Posicor	Roche Laboratories	Blood pressure, cardiovascular	6/20/1997	6/8/1998	1 year	Drug interaction, lowered heart rate in women	Not available
Bromfenac	Duract	Wyeth–Ayerst	Pain reliever	7/15/1997	6/22/1998	1 year	Liver failure	$89.7 million
Terfenadine	Seldane, Seldane-D	Hoechst Marion Roussel	Antihistamine	5/8/1985	2/27/1998	2.75 years	Heart rhythm abnormalities	Not available
Fenfluramine hydrochloride	Pondimin	Wyeth–Ayerst	Appetite suppressant	6/14/1973	9/15/1997	23.25 years	Valvular heart disease	Not available
Dexfenfluramine hydrochloride	Redux	Wyeth–Ayerst	Appetite suppressant	4/29/1996	9/15/1997	1.42 years	Valvular heart disease	$255.3 million

Source: From MedWatch Safety Reports and other online Internet records.

or obscure onsets.[5] These data complicate the causal relationship. Often a drug has been on the market for years before its ADRs become known. In a study of the *Physicians' Desk Reference* (*PDR*) black box warnings,[27] only half of the serious ADRs were detected and documented within 7 years of the drug's approval. Half of all drug withdrawals occur within 2 years of approval.[27,28] As mentioned, premarketing drug trials frequently miss ADRs because of their limited sample size.[27,29,30] See Table 28.3 for a list of prescription drugs withdrawn since 1997.

28.4.3.1 Incomplete Testing

The FDA has been under fire from at least two different sides. Industry and certain patient groups argue that the agency takes far too long to approve new drugs, whereas public interest groups argue that the agency is not thorough enough in its reviews of new drugs. Before drugs can be sold in the United States, the manufacturers must apply to FDA for the right to test them in humans, submit their test results to the agency, and then apply for the right to sell the drugs.

Although most experts agree that clinical trials completed before approval cannot detect all side effects caused by a drug because of the limited population size of the studies,[21] in some cases, serious ADRs found in premarketing trials fail to prevent a drug's release.[31] According to Lasser and colleagues,[27] there were reports that alosetron hydrochloride (Lotronex, GlaxoWellcome) was associated with ischemic colitis before its approval and subsequent withdrawal. In another example, David Willman of the *Los Angeles Times*[32] blamed new policies at FDA for the failure to prevent the marketing approval of grepafloxacin hydrochloride (Raxar, GlaxoWellcome, now GlaxoSmithKline), which was implicated in prolonging QT intervals (an electrocardiogram measurement), possibly leading to two deaths. Willman blamed the release and withdrawal since 1993 of seven drugs suspected in 1002 deaths on an FDA decision to "partner" with the pharmaceutical industry. After a 2-year investigation, the newspaper found "that the FDA approved each of those drugs while disregarding danger signs or blunt warnings from its own specialists."[32]

The *Los Angeles Times*[32] alleged that FDA often asks companies to conduct additional studies after a product is approved, if unresolved safety questions remain, but the "suggestion" is often ignored. According to research by Public Citizen,[33] in only 13% (11 of 88) of the drugs approved contingent on postmarketing studies did the companies who made those commitments actually complete the tests. Although FDA has the authority to withdraw drugs from companies that have not performed required postmarketing research, by April 2000, FDA had not withdrawn any drug because of a company's failure to complete a postapproval safety study. According to Willman,[32] FDA officials conceded that they do not know how often the studies are performed.

Of the seven drugs cited in the Willman research[32] that were approved then withdrawn from market, none were needed as a life-saving therapy, and all had alternatives that could have been substituted: alosetron hydrochloride (Lotronex; GlaxoSmithKline, Research Triangle Park, NC) was for irritable bowel syndrome, dexfenfluramine (Redux; Wyeth–Ayerst, Philadelphia, PA) was a diet pill, grepafloxacin (Raxar; Glaxo Wellcome, now GlaxoSmithKline, Research Triangle Park, NC) was an antibiotic, mibefradil (Posicor, Roche Laboratories, Nutley, NJ) treated high blood pressure, bromfenac (Duract; Wyeth–Ayerst, Philadelphia, PA) was a pain killer, Troglitazone (Rezulin; Warner–Lambert, now Pfizer NY, NY) was for diabetes, and cisapride (Propulsid; Janseen Cilag, Mignon, Italy) treated heartburn.[32] The recent withdrawals of Vioxx and Bextra, Cox-2 specific NSAIDs, are also examples of medications that were used to treat non-life-threatening conditions, were probably overprescribed (beyond any gastro-protective effect), and now that

they are no longer in the market, other medications have readily replaced them. As Lasser and collegues[27] suggest, "Some drugs represent a significant advance over existing drugs in the reduction of morbidity and mortality and warrant use despite limited experience. However, the drugs that do not represent a significant advance should be considered second-line drugs until their safety profile is better known." But patent life, blockbuster swings, investor, and stockholder demands, direct-to-consumer advertising, and prescribing habits mean that new medicines are often rushed to market.[27,34–40]

28.4.3.2 Inadequate Reporting

Currently, the United States has a voluntary, spontaneous reporting system called MedWatch. Healthcare professionals and drug companies report ADRs to the FDA using the MedWatch Drug Experience Report Form. Since 1962, drug manufacturers have been required to report all ADRs brought to their attention, and they file the majority of the 200,000+ ADRs reported annually. The MedWatch program allows health professionals and the public to voluntarily report serious ADRs, product quality problems, and medication errors for all FDA-regulated medical products by mail, fax, telephone, or over the Internet. Direct reports, primarily from healthcare professionals, increased 51% between 1998 and 2003. The FDA is interested in receiving reports of serious, new reactions associated with the use of drugs and biologic products used in the course of medical practice. In addition, there is interest in reactions to new drugs during their first 3 years of marketing in the United States. The FDA database currently contains over 400,000 reports, including any reports that indicate a lack of therapeutic response. The FDA does not collect reports of inappropriate use, prescriber errors, or administration errors.

The spontaneous reporting system of the United States has many good features. The input is global — covering the entire population of patients treated with a drug. It includes drugs and physicians in "real-life" situations of drug use, unlike the controlled setting of clinical drug trials. The system is inexpensive, requires a minimum of time, and does not interfere with the practice of the physician. Unfortunately, the overall effectiveness of the program has been found to be wanting.[41] Delays in sounding an alert can be a function of the need to verify and initiate any regulatory action. Alerting others usually occurs through published anecdotal reports. Leading a list of the limitations of MedWatch is underreporting — few physicians fill out the report forms. Although the MedWatch system is valuable for generating "signals" and forming hypotheses for epidemiological studies, it has been estimated that 95% of ADRs go unreported.[42] According to Schiff,[43] the primary reasons that doctors fail to report include complacency, fear, guilt, ambition, ignorance, diffidence, and indifference.

28.4.3.3 Inadequate Warnings

ADRs are believed to be one of the leading causes of death in the United States.[26,44] Patient exposure from new drugs with unknown toxic effects may be extensive. For instance, nearly 20 million patients in the United States took at least one of the five drugs withdrawn from the market between September 1997 and September 1998.[45] Three of the five drugs were new, having been on the market for less than 2 years.

"Boxed warnings" on medicine labels are described in the *Code of Federal Regulations*[46] as follows:

> Special problems, particularly those that may lead to death or serious injury, may be required by the Food and Drug Administration to be placed in a prominently displayed box. The boxed warning ordinarily shall be based on clinical data, but serious animal toxicity may also be the basis of a boxed warning in the absence of clinical data.

As this description lacks adequate criteria on when boxed warnings should be required, missing and inconsistent warnings are one of the causes for the high incidence of ADRs. For instance, the antipsychotic drug ziprasidone (Geodon; Pfizer), linked to heart toxicity, was not required to have a black box warning even though such a warning was required on six other drugs with prolonged QTc intervals. Sidney Wolfe[47] writes: "Although this is a dangerous inconsistency, it is somewhat predictable given the lack of clear FDA criteria for deciding on when a black box warning is necessary."

In the *PDR* study, Lasser and collegues[27] found that between 1975 and 1999, 548 new chemical entities (NCEs) were approved, and 56 (10.2%) acquired a new black box warning or were withdrawn from the market. The authors found, using a Kaplan–Meier survival analysis (see Table 18.1), that there was a 20% probability of a new drug acquiring black box warnings or being withdrawn from the market within 25 years and a 4% probability of a new drug being withdrawn from the market, half of which occurred during the first 2 years the drug was marketed. The authors also noted inconsistencies in the *PDR* safety warnings. For instance, four beta-blockers contained black box warnings about the dangers of abruptly discontinuing the drugs (which can exacerbate coronary artery disease), but three other beta-blockers had no such warning. They also found asynchronous dates on the warnings for drugs from the same class.

A relationship between inadequate warnings and liability should be clear. Since physicians rely on manufacturers' warnings, if a physician is sued for an ADR in their patient, and it is discovered that information related to precautions and warnings are deficient, the plaintiff or injured party, and the physician may in turn file litigation against a manufacturer. Indeed, in 20 years of consulting work in litigation, this author has observed that the overwhelming product liability claim against a pharmaceutical manufacturer is based on an inadequate warning, which is claimed to render the product defective. Since FDA regulations provide for manufacturers to strengthen their warning based on a reasonable association of risk, the manufacturer cannot (but frequently tries to) claim that the warnings were approved by the FDA and cannot be changed without FDA approval. The specific regulation is as follows:

> 21 USC 201.57 (e) Warnings. Under this section of the section heading, the labeling shall describe serious adverse reactions and potential safety hazards, limitations in use imposed by them, and steps that should be taken if they occur. The labeling shall be revised to include a warning as soon as there is reasonable evidence of an association of a serious hazard with a drug; a causal relationship need not have been proved.
>
> 21 USC 314.70 (c) Supplements for changes (in labeling) that may be made before FDA approval "Special Supplement-Changes Being Effected."
>
> . . . (2) changes labeling to accomplish any of the following:
>
> (i) to add or strengthen a contraindication, warning, precaution, or adverse reaction;
>
> (ii) to add or strengthen a statement about drug abuse, dependence, or overdosage; or
>
> (iii) to add or strengthen an instruction about dosage and administration that is intended to increase the safe use of the product;
>
> (iv) to delete false, misleading, or unsupported indications for use or claims for effectiveness.

28.4.3.4 Investigator Fraud

Another problem in detecting and reporting ADRs has been the discovery of numerous instances of proven investigator fraud, which can vary from bewildered ignorance to deliberate dishonesty:

- Some fail to document and report what they should clearly recognize as an ADR.
- Some are so "invested" in the study drug, believing so strongly in its safety, that their enthusiasm overshadows sound scientific and regulatory standards.
- Some choose not to report ADRs — either because it requires too much effort or because they are deliberately attempting to defraud.
- Some are confused over what constitutes an ADR.
- Many ADR reports lack essential information.
- Most are weak in causality assessment (whether the drug actually caused the ADR).

28.4.3.5 Quality Control, Safety, Manufacturing

A recent book, (*Risky Business: Managing The Quality of America's Medicines*, Robert A. Rhodes, FDA News, www.fdanews.com) attracted my attention, and I quickly read it as this chapter and the book was being readied for submission to the publisher. Rhodes, a veteran of 25 years in Quality Control in the Industry and now a member of the Weinberg Group, recommends the book to anyone thinking about becoming a quality professional in an FDA-regulated industry, as a means of achieving a sufficient picture of the complexity and intricacy of what it is like, through the eyes of someone who has "been there, done that." To me the book reads like a "what ever could go wrong, will go wrong," and quality systems in manufacturing and distribution are established and necessary to help protect the integrity of the product and the safety of the patient, which will ultimately help protect the Company, which could suffer serious financial and reputation setbacks for a defective product. As stated elsewhere, while most drug liability litigation flows out of inadequate warnings, from time-to-time, ineffective manufacturing controls and procedures are to be blamed for product failure, patient injury, and litigation and regulatory actions against the pharmaceutical manufacturer. I strongly recommend this book to anyone working in the pharmaceutical industry, not just those involved in "Quality."

28.4.4 Types of Drugs Most Frequently Involved in Litigation

Although almost any drug can cause an adverse reaction of varying severity in at least some individuals, a number of drugs can be singled out as causing approximately 90% of reported drug reactions: adrenal steroids, aminoglycosides, anticoagulants, antimicrobials, antineoplastics, aspirin, bronchodilators, digitalis, diuretics, insulin, and NSAIDs. These agents have commonly been cited for the past 20 years. With the exception of digitalis, the use of which is declining, the "usual suspects" continue to be drawn from the same drug classes. Certain drug categories routinely appear in ADR studies.[48] Examples are:

- Anticoagulants (heparin, warfarin)
- Antimicrobials (penicillins, cephalosporins, sulfonamides)
- Cardiac agents (digoxin, diuretics, hypotensives, quinidine)
- Central nervous system agents (analgesics, anticonvulsants, sedative-hypnotics)
- Diagnostic agents (radiocontrast media)
- Hormones (corticosteroids, insulin)

As a rule, it has been this writer's observation that the longer the drug is on the market, the more we learn about the drug, and the more information is in the package insert, the less the drug company or manufacturer faces liability for inadequate warnings. The new drugs are usually the subjects of product liability suits.

The PIAA studies[20,24] found that the most frequent drug classes and errors involved in its claims were antibiotics (failure to note documented allergy, most appropriate drug not used, drug inappropriate for medical condition); glucocorticoids (incorrect dosage, communication failure between physician and patient, failure to monitor for drug side effects); and narcotic and non-narcotic analgesics or narcotic antagonists (drug inappropriate for medical condition, incorrect dosage, failure to monitor for drug side effects, drug dependence).[49] The following sections describe drugs of note — those that have been the subject of large or highly publicized lawsuits.

28.4.4.1 Vaccines

Although vaccines are relatively safe, they have caused problems in the past, and any vaccine can cause an ADR in a "predisposed individual."[5] As mentioned, the government has an interest in controlling disease outbreaks, so vaccines are routinely given to millions of people within months of each other. With that many people receiving a vaccine, even very low incidence of an ADR can still result in high rates of litigation, and as children receive a large number of vaccines (often several vaccines at once), jury awards can be high (see Section 28.2.1).

The production of vaccines has marched hand-in-hand with production mistakes, medication errors, and ADRs. For example, in October 1954, the National Foundation for Infantile Paralysis, which had paid for the development of the Salk vaccine to prevent polio, announced it was ordering enough of the drug to vaccinate 9 million children — without waiting for the results of a field trial.[50] This became the famous "Cutter Incident," when the doctor who conducted the field trials stated the vaccine was 60 to 90% effective, assuming that it was ineffective for the children in the trial who developed polio. Newspapers, parades, air-raid sirens, and erroneous statements all helped create the euphoria over the treatment. After vaccine shortages, conspiracy charges, and somewhat unregulated manufacturing to increase production, postinnoculation polio was seen in children inoculated by vaccine from the Cutter Laboratories (Berkeley, CA, U.S.A.) which turned out to have live polio virus in some of its batches.

In 1976, another vaccine case illustrated how dangerous a mistake involving vaccines can be. Between October 1, 1976 and December 14, 1976, more than 40 million people were vaccinated against swine flu (a virus similar to the 1918 Spanish flu virus that killed so many during World War I). The feared epidemic never manifested, but there was a tenfold increase in GBS (in which the body's immune system attacks its peripheral nerves), which thousands of people contracted.

Numerous other vaccines have been linked to a variety of MedWatch — The FDA Safety Information and Adverse Event Reporting Program FDA has requested that sponsors of all NSAIDs make labeling changes to their products. The FDA recommended proposed labeling for both the prescription and OTC NSAIDs and a medication guide for the entire class of prescription products. All sponsors of marketed prescription NSAIDs, including Celebrex (celecoxib), a COX-2 selective NSAID, have been asked to revise the labeling (package insert) for their products to include a boxed warning, highlighting the potential for increased risk of cardiovascular (CV) events and the well described, serious, potential life-threatening gastrointestinal (GI) bleeding associated with their use. The FDA regulation 21CFR 208 requires a Medication Guide to be provided with each prescription that is dispensed for products that FDA determines pose a serious and significant public health

concern. Read the complete MedWatch 2005 Safety summary, including a link to the updated Drug Information Page and Medication Guide at http://www.fda.gov/medwatch/SAFETY/2005/safety05.htm#NSAID. Autoimmune disorders, such as encephalitis, lupus-like reactions, arthritis syndromes, arthorpathies, and transverse myelitis.[5] Other vaccines have been linked to ADRs:

- *Rotavirus* — the most common cause of severe diarrhea in infants, with approximately 125 million cases worldwide per year and 600,000 deaths — vaccines have been linked with intussusceptions (a problem with the intestine in which one portion of the bowel slides into the next) at a rate of between 1 in 5000 and 1 in 11,000 infants.[51] The only rotavirus vaccine approved in the United States, RotaShield (Wyeth–Ayerst), was withdrawn from market on October 22, 1999, about 1 year after licensure.
- The *diphtheria, pertussis, and tetanus* (*DPT*) vaccine, which has prevented more than 95% morbidity from these diseases, has been linked with convulsions, encephalitis, and sudden infant death syndrome.[52]
- The *hepatitis B* vaccine has been linked to anaphylaxis.
- The *measles* vaccine has been linked to thrombocytopenia and anaphylaxis.
- *Tetanus* toxoid-containing vaccines to GBS, brachial neuritis, and anaphylaxis.

What is, perhaps, most disturbing is the assertion by Coulter and Fisher[53] in 1985 that the ADRs from DPT should not have happened. They document that Japan switched to an acellular form of the DPT virus in 1981, which was just as effective in preventing the diseases but had fewer ADRs. Shoemaker[5] reports that "In Japan, the Ministry of Health, instead of trying to cover up problems with the vaccines, chose to find a solution." According to Shoemaker, it took almost 20 years for the United States to stop using the whole-cell version of the vaccine, and manufacturers are still distributing the whole-cell version in third-world countries "undoubtedly because it is cheaper to make."[5]

Vaccines are currently in development for allergies, anticholesterol, behavioral addictions, cancer, *Candida albicans*, cat allergy, *Chlamydia trachomatis*, cytomegalovirus, *Escherichia coli*, genital herpes, gonorrhea, *Helicobacter pylori*, hepatitis C, hepatitis E, herpes simplex, juvenile diabetes, *Listeria monocytogenes*, malaria, multiple sclerosis, nicotine, peanut allergies, periodontal disease, ragweed, respiratory syncytial virus (RSV), rheumatoid arthritis, ringworm, *Staphylococcus aureus*, *Streptococcus* genus, syphilis, and tuberculosis.

28.4.4.2 Fen–Phen and Other Diet Drugs

The use of drugs to treat obesity has had limited success because of the danger of unintended side effects. Some of these side effects are, at worst, a nuisance, but some of them can be life threatening. Compounding the risk of taking diet drugs as directed is by taking them in doses exceeding the recommended amount or in an "off-label" manner. Using larger doses, extending the duration of treatment, or combining two or more drugs to enhance the same effect are examples of this practice. The Fen–Phen diet is an example of combining two drugs for assistance in achieving weight loss. Each had been tested and approved as safe and effective for short-term use in morbidly obese subjects. The combined therapy proved to be so well received that millions of prescriptions were written for two drugs that were not approved or tested for safety in combined use. In addition, the duration of therapy was arbitrarily extended from the approved 6 to 8 weeks (short term), to more than 1 year. The result was that a population experiment in which a potentially deadly adverse effect, PPH, was discovered after millions of doses had been prescribed and administered.

PPH was a problem in Europe in the 1960s for an appetite suppressant, Aminorex. There is no method for screening potential dexfenfluramine patients for susceptibility to PPH: it is a silent killer, with no early symptoms, affects predominately women in their early 30s and 40s, and that there is no cure.[54] The risk of PPH among anorexic agents is significantly elevated, but the absolute incidence is still small: 28 cases per million person-years of exposure, comparable to the fatality risk from penicillin-caused anaphylaxis. The risk of death from untreated obesity is perhaps 20 times higher than the estimated mortality from PPH among patients given appetite-suppressant drugs.[55]

The Chairman of FDA's Advisory Committee commented on the risk of death from developing PPH. "We have had what I think appears to be a reasonable estimate of the risk of deaths from pulmonary hypertension. We need to understand clearly that if a million patients take this drug, at least a couple dozen of them will die annually as a result of this complication. That seems the best estimate. This is something that has to be weighed seriously." The appearance of heart valvulopathies in otherwise asymptomatic people in their thirties or forties was unexpected and caught patients and practitioners by surprise. Both fenfluramine and phentermine cause an increase in the amount of serotonin available in the body, which can cause cardiac valvulopathies.

On September 15, 1997, FDA asked the manufacturers of dexfenfluramine (Redux; manufactured for Interneuron Pharmaceuticals by Wyeth–Ayerst) and fenfluramine (Pondimin; Wyeth–Ayerst) to voluntarily withdraw both treatments from the market because of findings that indicate approximately 30% of patients taking the combined drugs had abnormal echocardiograms, even if they had no symptoms. Both companies agreed. FDA is not requesting the withdrawal of phentermine, the third widely used medication for obesity.

Additional ADRs linked to diet pills include psychosis;[56] myocardial ischemia;[57] drug interactions, such as the interaction of fenfluramine with imipramine, fenfluramine with amitriptyline or desipramine, or the toxic reaction between fluoxetine and phentermine; and the release of serotonin while inhibiting its reuptake,[58] contributing to hyperserotonin reactions. When the next craze takes hold of patients and their physicians, hopefully physicians and pharmacists will take a more vocal position and recommend restraint, until some proof of efficacy and lack of toxicity is shown for new faddish off-label combinations.

28.4.4.3 Nonsteroidal Anti-Inflammatories

The efficacy of NSAIDs in the treatment of a wide variety of disorders is well established. An estimated 1.2% of the U.S. population takes NSAIDs regularly, and more people take them intermittently. NSAIDs have also been involved in the prescription-to-OTC shift (Motrin to Advil; Naprosyn to Aleve) and so more — and more potent — NSAIDs are being used. The NSAIDs, according to FDA, account for the largest number of ADR reports.

Gastrointestinal effects are the most frequently reported ADRs associated with NSAIDs and include nausea, vomiting, dyspepsia, and diarrhea. Gastrointestinal bleeding, ulceration of the GI tract, and renal toxicity in predisposed patients are also common. The NSAIDs are also linked to cholestatic hepatitis and GI hemorrhage, which may be fatal. Salt and water retention with edema is a well-known effect of NSAIDs. Hyperkalemia and reduced excretion of potassium have also been reported. Most NSAIDs can inhibit platelet aggregation and prolong bleeding time. Elderly patients excrete the drug more slowly, putting them at greater risk of bleeding.

The NSAIDs have been increasingly incriminated in chronic peptic ulcerations. Patients on short- and long-term therapy should be instructed on the signs of GI bleeding and ulcer perforation, and periodic monitoring of renal function is advisable. In "*The Seven Pillars of Foolishness*," Dukes,[59] gives an account of some past disasters occurring from the use of NSAIDs class. One of the NSAISs Dukes discusses is Benoxaprofen, introduced in the

early 1980s. A discussion of the Benoxaprofen problem will be presented to give a historical focus to some recent problems with the NSAIDs. Four NSAIDs that have caused concern recently include rofecoxib, valdecoxib, benoxaprofen, and ketorolac.

Rofecoxib, a Cox-2 specific inhibitor NSAID, is one of the drugs most recently withdrawn from market (Table 28.3). Rofecoxib (Vioxx; Merck & Co.) was voluntarily withdrawn from the U.S. and worldwide markets on September 30, 2004, because of an increased risk of heart attacks and strokes confirmed during investigations to determine if the Rofecoxib was effective in preventing the recurrence of colon polyps. A postwithdrawal Advisory Committee questioned the need to withdraw the drug, and the Company has released statements suggesting that Vioxx may be re-introduced to the market. This COX-2 selective NSAID has been on the market for more than 5 years for treating osteoarthritis, rheumatoid arthritis, acute pain, and menstrual symptoms. On approval, it had gone through an expedited (6 months) "priority review" because the drug "potentially provided a significant therapeutic advantage over existing approved drugs due to fewer gastrointestinal side effects, including bleeding," according to FDA.

The original safety database on Vioxx included approximately 5000 patients on rofecoxib and did not show an increased risk of heart attack or stroke. A June 2000 study, VIGOR (VIOXX GI Outcomes Research), was primarily designed to look at the effects of rofecoxib on ADRs such as stomach ulcers and bleeding. The study showed that patients taking rofecoxib had fewer of the side effects than patients taking naproxen. The study also showed a greater number of heart attacks in patients taking rofecoxib. The FDA's response to the VIGOR study was to include new safety information on the Vioxx label in April 2002. Merck then began to conduct longer-term trials to obtain more data on the risk for heart attack and stroke with chronic use of Vioxx. Additional information on Vioxx is available at www.fda.gov/cder/drug/infopage/vioxx.

Valdecoxib (Bextra; Pfizer), a second COX-2 specific inhibitor NSAID was taken off the market a few months after Vioxx. In addition to a higher risk of serious skin toxicity, patients given Vioxx after Coronary Bypass Surgery had a higher rate of myocardial infarctions.

In the aftermath of the Vioxx and Bextra withdrawals, the FDA studied the CV and GI risks in great detail. A very recent request for strengthening the warnings and precautions on all NSAID products was released, and reads as follows:

> MedWatch — The FDA Safety Information and Adverse Event Reporting Program
>
> FDA has requested that sponsors of all non-steroidal anti-inflammatory drugs (NSAID) make labeling changes to their products. FDA recommended proposed labeling for both the prescription and over-the-counter (OTC) NSAIDs and a medication guide for the entire class of prescription products. All sponsors of marketed prescription NSAIDs, including Celebrex (celecoxib), a COX-2 selective NSAID, have been asked to revise the labeling (package insert) for their products to include a boxed warning, highlighting the potential for increased risk of cardiovascular (CV) events and the well described, serious, potential life-threatening gastrointestinal (GI) bleeding associated with their use. FDA regulation 21CFR 208 requires a Medication Guide to be provided with each prescription that is dispensed for products that FDA determines pose a serious and significant public health concern.

Read the complete MedWatch 2005 Safety summary, including a link to the updated Drug Information Page and Medication Guide at http://www.fda.gov/medwatch/SAFETY/2005/safety05.htm#NSAID

Benoxaprofen was an antirheumatic drug compound that early clinical studies suggested might relieve gastric problems associated with the class. This agent was approved for marketing in the early 1980s and touted as a new miracle drug, thought by some to improve the arthritis disease, in addition to serving its proven efficacy of acting as an anti-inflammatory.

When the real problems with benoxaprofen emerged, however, they were more serious: It was apparently killing elderly patients with hepatic disorders, inducing massive photosensitivity, and causing oncholysis (separation of the nail plate from the bed) in about 15% of patients. It seems probable that at least 70 elderly patients died, and many more people suffered.[60] Shortly after its well-publicized entry into the U.S. market, the manufacturer of benoxaprofen voluntarily withdrew its product as it caused fatal cholestatic hepatitis.[60–63] This action immediately followed news of suspension of the license to sell benoxaprofen in the United Kingdom.[64]

Ketorolac, in both injection and tablet form, has had its safety questioned. Even short-term parenteral treatment with ketorolac can cause gastric ulceration. According to the manufacturer (Syntex) by December 1992, approximately 16 million patients worldwide had received all formulations of ketorolac.[65] By 1993, there had been a total of 923 reports of serious ADRs from ketorolac, 838 from the United States, including GI (n = 203), hematological (n = 181), renal (n = 124), hypersensitivity (n = 107), and neurological (n = 111) reactions. Fatal outcomes were reported in 97 cases worldwide by April 1993, and by that time, 26 million patients had received the drug. Death resulted from GI bleeding and perforation (n = 47), renal insufficiency (n = 20), anaphylaxis and asthma (n = 7), hemorrhagic reactions (n = 4), miscellaneous causes (n = 13), and unexplained (n = 6) reactions. These data have prompted different reactions in different countries, ranging from no action (except for updating the prescribing information sheet) in the United States to limiting its use in many countries to a request for withdrawal from market in Germany and France. Other NSAIDs that have been introduced into the market and removed for safety reasons include Zomax, Suprol, and Duract.

28.4.4.4 Corticosteroids

Corticosteroid therapy can be of great benefit; they are the strongest drugs available for reducing inflammation. They are used to treat arthritis, multiple sclerosis, chronic obstructive pulmonary disease, and certain emergencies (asthma attacks, anaphylaxis, and brain swelling, for instance). They can have phenomenal success in treating tough skin conditions, such as eczema and psoriasis. When inflammation is severe, corticosteroids can save lives, but adverse effects from their use appear frequently in the medical literature — as might be expected from a group of drugs that exert many pharmacological effects and are routinely used or tried experimentally for a variety of pathological conditions. The long-term use of corticosteroids is often limited by Cushingoid side effects, which include the well-known facial puffiness and weight gain, among other complications such as activation of tuberculosis, cataracts, diabetes, ecchymosis, hirsutism, hypertension, infections, obesity, osteoporosis, phlebitis, poor healing, and renal lithiasis.

Short-term use of glucocorticoids, even in massive dosages, is less likely to produce harmful reactions. They can, however, produce a variety of effects that are neither limited to high doses nor to long-term therapy. When a low dose of steroids (prednisone) was given for several months to a 38-year-old man to treat eczema of his hands and feet, he developed bilateral avascular necrosis (AVN) of the femur. He therefore had total bilateral hip replacement, and several experts have attributed his AVN to the steroid administration. The man sued his allergist, who settled the lawsuit shortly before trial for approximately $400,000. Most practitioners, however, are unaware of the risk of short-term or low-dose steroids. Yet many of these cases can be found in the courts, the literature, and the MedWatch databases.

Topical corticosteroids produce anti-inflammatory, anti-itching, and vasoconstricting effects. A wide variety is available (described in detail in *Drug Information 2004*[66] and other references and compendia). Such steroids are classified into five categories that reflect a

TABLE 28.4

Systemic Disorders That Can Occur with the Use of Corticosteroids

Disorders	Glossary
Cardiovascular disorders	Hypertension is one of the Cushingoid effects of corticosteroids.
Dermatological disorders	Various dermal manifestations have also been reported. Topical steroids may cause facial edema (moon-face) atrophy or thinning of the epidermis and dermal collagen, drying of the skin, telangiectasis, fragility of the skin blood vessels, purpura (easy bruising), and atrophic striae.
Endocrine and metabolic disorders	The exogenous administration of glucocorticoids can result in hypothalamic-pituitary-adrenal axis (HPA) suppression, which may subsequently lead to adrenal atrophy.[66,67] The degree of adrenal suppression is dependent on the dosage, duration, frequency, time, and route of administration of the specific glucocorticoids. At least one patient who received prednisone for neurological symptoms developed Cushing's syndrome.[68]
Genitourinary disorders	Steroid use has been linked to renal lithiasis.
Gastrointestinal disorders	Corticosteroid ADRs include the development, reactivation, perforation, hemorrhage, and delayed healing of peptic ulcers.
Immunological disorders	Delayed or poor wound healing, increased susceptibility to infections, masking of the symptoms of infections, thrush infections of the mouth and anaphylactoid reactions have been reported.
Musculoskeletal and connective tissue disorders	Steroid use may cause muscle wasting, pain and weakness, and atrophy of the protein matrix of bone, which can lead to osteoporosis, vertebral compression fractures, aseptic necrosis of femoral and humeral heads, and other pathologic fractures of the long bones. AVN has been reported following short courses (1 week). Local injections into a patient's joint (intra-articular) can produce systemic as well as local effects. Patients who have received a local injection aimed right into a joint should be given careful instructions not to overuse those joints, even after symptomatic relief of the pain. The repeated administration of intra-articular injections may result in damage to joint tissues. Septic arthritis — which inclu des pain, local swelling, further restriction of joint movement, fever, and malaise — is also possible. Because there is latency between the end of steroid therapy and the onset of symptoms of AVN, patients should be informed of the potential risk of osteonecrosis following the use of steroid medication. Complaints of hip pain in people who have previously been prescribed steroids should produce a high index of suspicion for underlying osteonecrosis of the femoral head.

Neurological disorders	Ischemic neuropathy, restlessness, and seizures are possible neurological ADRs.
Nutritional disorders	The primary nutritional ADR from corticosteroids is weight gain and obesity.
Obstetrical disorders	Small birth weight and necrotizing enterocolitis of a fetus resulted from the mother's use of a corticosteroid.
Ophthalmoligical disorders	Ocular complications following local or systemic administration of steroids include glaucoma, cataracts, adverse influence on specific ocular infections, pseudotumor cerebri, ptosis, mydriasis, subjective visual complaints, visual field defects, systemic absorption of the topical medication, conjunctival and palpebral petechiae, epithelial punctate keratitis, and, possibly, corneal and scleromalacia.[69]
Psychiatric disorders	Glucocorticoids are known to produce mental changes that range from moderate mood changes, severe depression, and euphoria to psychosis. Steroid psychosis is characterized by a delirium and a clouded sensorium. The onset of symptoms usually occurs within 5 to 30 days after therapy is initiated. The incidence rate appears to correlate closely with dosage level (a rate of 1.3% when the dose was 40 mg or less; 4.6% when it was between 41 and 80 mg; and 18.4% when it was 80 mg or more.
Withdrawal symptoms	The abrupt discontinuance of steroids in long-term therapy patients can result in withdrawal syndrome, even without evidence of adrenal suppression. To minimize morbidity associated with adrenal insufficiency, discontinuing corticosteroid therapy must be gradual. During withdrawal therapy, increased supplementation may be necessary during times of stress. Symptoms of adrenal insufficiency that result from too rapid withdrawal include nausea, fatigue, anorexia, dyspnea, hypotension, hypoglycemia, myalgia, and arthralgia. Continued supervision after therapy termination is essential because a sudden reappearance of a severe manifestation of the disease being treated can occur. Whether or not there is improvement in the disorder that is being treated, clinicians frequently attempt to withdraw patient from glucocorticoid therapy before such therapy has gone on for very long.[70]

steroid's vasoconstricting effects and its action on psoriasis. In general, an application of topical steroids does not demonstrate systemic effects. However, systemic ADRs can occur when the drugs are used on large skin areas, for prolonged periods of time, with occlusive dressings that prevent evaporation of moisture and drive the drug into the dermis, when more potent preparations are applied to areas of greater than average absorptive abilities (for example, scrotum, scalp, macerated skin), or when they are used on infants and children. In addition, recent evidence suggests that excessive or prolonged use of topical fluorinated steroids (for example, triamcinolone) during pregnancy can affect the intrauterine growth of the fetus.

Corticosteroids can produce a variety of devastating and systemic effects. Some of the ADRs associated with corticosteroids are listed in Table 28.4.

28.4.4.5 Noteworthy Class Actions: Rezulin, Baycol

Rezulin (troglitazone, Warner Lambert, now Pfizer) was approved in January 1997 to treat adult-onset diabetes and was hailed as a drug to treat patients who have failed other therapies (untreated diabetes can cause heart and kidney failure, blindness, and other problems). It was a "fast track" drug with an approval process that took only 6 months, despite objections from several FDA scientists and the death from liver failure of one of the study participants. The FDA received reports of liver failure and in December 1997, it was banned in Great Britain. The FDA, at that time, however, only ordered stronger liver toxicity warnings on the drug label — the warning was strengthened four times between 1997 and June 1999 — even though several patients who had monthly liver tests still experienced sudden liver failure and died.[67] In March 1999, an FDA epidemiologist warned that "Rezulin was among the most dangerous drugs on the American market," and that "patient monitoring would not protect them from liver failure."[67] An advisory committee recommended the drug be made available only to patients whose diabetes was not well controlled by other drugs. In March 2000, an FDA official Robert Misbin wrote to Congress, "I am writing to enlist your aid in convincing my superiors at FDA that Rezulin should be removed from the market because of its unacceptably high risk of causing liver failure."[68]

Even the efficacy of troglitazone was in suspect. After the drug was removed from the market, one FDA official wrote that it had been approved and kept on the market so long because it was "shown to reduce or delay long-term serious effects of diabetes, including death." But when asked about the basis of that claim, an FDA spokesperson said "those findings were not intended as definitive scientific observations."[67] With the help (and possible conflicts of interest) of FDA officials and the physicians carrying out a National Institutes of Health (NIH) study,[69] the drug manufacturer was able to keep FDA from banning troglitazone for 27 months.[70] The drug was eventually withdrawn from the market in the United States on March 21, 2000, but keeping the drug on the market as long as possible was profitable. At its peak sales in January 1999, 488,000 prescriptions were filled, and during its 3 years on the U.S. market, troglitazone generated $2.1 billion for Warner Lambert.

The withdrawal of troglitazone came only after a whistleblower shared his findings and internal email with a *Los Angeles Times* reporter. A series of articles by David Willman,[32,69,71] raised questions about irregularities and conflicts of interest in the study and approval of the drug. Litigation over troglitazone is still ongoing. In 2003, an appeals court overturned a lower court, which had denied class action status to Rezulin cases. In January 2004, Pfizer set aside $975 million to cover 35,000 settled or withdrawn injury claims. In March 2004, a federal grand jury requested testimony from former Warner–Lambert employees. In April 2004, a jury awarded $2 million as compensatory damages to a woman. In June 2004, a Los Angeles jury found that facts about the drug did not support responsibility for the death of two patients and the injury of another. In July 2004, an Illinois class action

lawsuit was settled for $60 million, and an $11.55 million award was upheld for a man who died 1 month after starting the drug.[72]

28.4.4.6 Drug-Induced Hepatotoxicity

Of course, Rezulin is not the only drug implicated in liver toxicity. In fact, the FDA recently posted a notice on its website (http://fda.gov/), describing a special interest and monitoring of drug-induced hepatotoxicity:

> Drug-Induced Liver Toxicity
>
> > Momentum and interest continue to grow concerning the rising incidence of liver toxicity, uncommon but serious, caused by prescription drugs, over-the-counter medications, or dietary supplements that are often combined with special diets and alcohol consumption, in addition to environmental chemicals. The liver is a most marvelous organ that usually protects us against injury from foreign substances, and is very robust in its capacity to withstand damage and heal itself. In a few people, the ability to resist and heal is not adequate, or the injury is so great that serious liver damage results, with progression to acute failure, and to death or transplantation. Drug-induced liver injury has become the most frequent cause of acute liver failure in the United States, exceeding all other causes combined. Drug-induced liver injury also remains the major single reason for regulatory actions concerning drugs, including failure to approve, withdrawal from the market, restrictions on use, and warnings to physicians. (*emphasis added*).
>
> > This web site is sponsored by the Hepatotoxicity Steering Committee, (HepToxSC) jointly made up of interested persons from the Food and Drug Administration Center for Drug Evaluation and Research (FDA/CDER), the Pharmaceutical Research and Manufacturers of America (PhRMA), and the American Association for the Study of Liver Diseases (AASLD). Members of this regulatory-industry-academic research group have been meeting annually for several years since the first public conference in Chantilly VA in February 2001, with additional quarterly telephone conferences. Material from the most recent annual meeting, January 28, 2005, is shown here; the older material is available in the background by clicking on the underlined listings.
>
> > We call attention also to the web site for the Drug-Induced Liver Injury Network (DILIN) sponsored since 2003 by the National Institutes of Health (NIH), Liver Disease Branch of the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK).
>
> > We encourage you to contribute your ideas and comments about this topic. Please send your thoughts to John R. Senior, M.D. (john.senior@fda.gov) or Lana Pauls, M.P.H. (lana.pauls@fda.gov).

The reader should note the emphasis in the above paragraphs: drug-induced hepatotoxicity is the leading cause of acute liver failure and the number one reason for regulatory actions against drugs in the United States!

Baycol (cerivastatin, sold as Lipobay in Europe, Bayer) is a statin, a class of cholesterol-lower drugs. Statins are the most prescribed drugs in the United States, with more than 12 million people taking them, and more than 700,000 people in the United States taking cerivastatin.[73] It received marketing approval on June 26, 1997 and was voluntarily removed from the market on August 8, 2001 because of its link to 100 deaths and several injuries from potentially the muscle disease rhabdomyolysis. The other statins — lovastatin (Mevacor; Merck); pravastatin (Pravachol; Bristol-Myers Squibb); simvastatin (Zocor; Merck); fluvastatin (Lescol; Novartis); atorvastatin (Lipitor; Parke-Davis); rosuvastatin

(Crestor; Astra-Zeneca); and lovastatin + niacin (Advicor; Kos Pharmaceutical) — that can also cause rhabdomyolysis remain on the market. Although scientists agree that the other statins "seem to have essentially identical safety profiles and benefit–risk ratios,"[77] FDA said the ADRs associated with Baycol "have been reported significantly more frequently than for other approved statins."[74]

When initially recommending cerivastatin for approval, after testing in 3343 patients from the United States, Japan, and China, an FDA official cited rhabdomyolysis as one of four potentially serious ADRs. After approval, a warning was added in 1999, and the labeling was changed again in June 2001. Just 2 months later, Baycol was withdrawn from the market. In 2003, the first Baycol lawsuit went on trial in Texas and involved a victim who did not die or require kidney transplant;[74] the verdict favored Bayer. The primary legal complaint in subsequent lawsuits is that the manufacturer failed to adequately warn of the known dangers and complications of the drug. "In clinical trials, there is a "filtering" of adverse events in drug-treated groups in that if the principal investigator expresses the opinion that the adverse event was not related to the drug, one researcher says, "the adverse event is excluded from analysis."[74] Kauffman,[75] discussing Lipitor, suggests that the cost–benefit analysis of statins is another possible area of litigation, saying, "The absolute risk reduction is … one in 667 per year. The cost of a month's supply of Lipitor at 40 mg/day is $1,500/year, reflecting a cost of $1 million to prevent one death among 667 people taking the drug for one year."[75]

In March 2004, Bayer settled with its insurance carriers, eliminating the insurers rights of litigation and settlement and setting the limits of liability at $1.2 billion. In July 2004, the company settled 2771 cases for $1.06 billion. In September 2004, Bayer settled another 2861 product liability cases for $1.09 billion. There are still 7577 such suits pending.[4]

28.5 Market Entry and Subsequent Withdrawal

In 1997, 39 new drugs were approved by the FDA. Now, five of them have been taken off the market (see Table 28.4) and an additional two have new boxed warnings. Thus, seven drugs approved that year (18%) have already been withdrawn or had a black box warning 4 years after approval. According to one study, 20% of drugs will be withdrawn or have a black box warning within 25 years of coming on the market.[27]

In 10 years (1988 to 1998) the number of drugs approved first-in-the-world by the FDA jumped from 4 to 66%. The FDA approved 80% of the new drug applications at the end of the 1990s, but only 60% at the beginning of the decade. The FDA has recently been slow in removing drugs that have already been withdrawn in Europe. There are increasing reports of drug approvals after FDA scientists and drug advisory committees have recommended against approval. Charges have been made that researchers and advisory committee members deciding on a drug's safety and approval often have conflicts of interests. Several medical researchers have concluded that new drugs should only be given with extreme caution. The FDA argues that many good therapies are lost as physicians fail to read and heed the drug warnings it issues. There are many reasons that drugs might be approved and later withdrawn, and there are many people and groups pointing fingers at others. A New Drug Safety Board has been appointed, constituted by members primarily from the agency, and Congress has substantially increased the funding for FDA safety monitoring.

Many drugs are simply discontinued for reasons that may not have to do with safety. The OTC drugs and drug products used as a compound in other drugs may also be withdrawn.

So occasionally different numbers of withdrawn drugs are reported. Table 28.3 lists prescription drugs withdrawn from market for safety reasons between 1997 and 2004.

28.6 Deep Pockets and Efforts to Limit Litigation

The Bush administration is now asking the courts to decide whether consumers can recover damages for ADRs. In what it admits is a policy shift, the Department of Justice now suggests that consumers cannot recover damages as the drugs involved were approved by FDA. The administration contends that allowing such lawsuits undermines public health and interferes with the federal regulation of drugs and devices by allowing juries and judges to second-guess FDA experts.[18] It also argues that if such lawsuits succeed, good products might be removed from the market. The federal court of appeals in Philadelphia agreed with the administration in the case of a defective heart pump case.[76] In *Horn v Thoratec Corp*. the company argued that the plaintiff's claims were preempted by the Medical Device Amendments (MDA) to the *Food, Drug, and Cosmetic Act*. The Third Circuit agreed and affirmed the district court opinion. Five other circuit courts have found that personal injury suits are improper for FDA-approved medical devices. One circuit court found to the contrary.[76] Although the Supreme Court weighed in on this issue in 2001,[77] that decision was based on the Court's finding that Congress had vested enforcement authority solely to the FDA in the MDA changes to the FD&C law. The MDA preemptions will probably be revisited by the Supreme Court, which previously ruled that FDA regulation does not preempt state law or local regulation.[78] These changes have only been applied to medical devices, not to drugs, at this time.

In 2002, the administration shifted the federal government's previous policy by making FDA a party in product liability lawsuits. The administration believes that consumer lawsuits are a burden on the economy and increase the cost of healthcare. Senator John Edwards, who was a plaintiffs' attorney, disagrees, and that is why he helped draft a bill defining patients' rights, including the right to sue. Although the bill passed the Senate (S-283), threats of a veto stalled it. A Bush-Norwood compromise was passed in the House that would allow patients to sue their insurers in state court but pain-and-suffering and punitive damages would be capped at $1.5 million each. The bill, however, never became law.[79]

Canada limits lawsuits against drug manufacturers. According to one estimate, one third to one half of the drug price difference between the United States and Canada is a result of the reduced litigation in Canada.[80] Maurice D. Hinchey, the New York representative to Congress, however, believes that the Bush administration has gone too far, saying it has "taken the FDA in a radical new direction, seeking to protect drug companies instead of the public." He recently persuaded the House to cut $500,000 from the FDA chief counsel budget to slow its "aggressive opposition to consumer lawsuits."[81]

28.7 Summary and Conclusions

Clearly, ADRs continue to induce injury and harm to patients, despite careful prescribing, dispensing, and administering. New drugs represent unknown risks and require intense scrutiny and monitoring in their early-market stages. Poor monitoring, selecting more toxic drugs, or taking poor patient histories increases a patient's risk of partially reversible morbidity and mortality. Careful monitoring by pharmaceutical manufacturers during

clinical trials, early post marketing use, and throughout the life of the product, with labeling updates and promulgated adequate warnings to prescribers and patients will improve the safety of medicines and limit the liability of the pharmaceutical manufacturers.

The best science in the world, as evidenced and described in this book, has to be balanced and tempered with thorough safety monitoring in order to safely provide the new drug discoveries with patients.

References

1. Fitzgerald, G., PhD, former FDA officer in the Center for Drug Evaluation and Research, Personal communication, 1992.
2. Dukes, M.N.G., Risk and safety in medicine, in *Meyler's Side Effects of Drugs*, Elsevier, Amsterdam, 1996.
3. Merli, R., Lawsuits against drug companies put a premium on risk management, *Pharm. Insider*, 2004.
4. American College of Clinical Pharmacology, Bayer settles 2,861 more Baycol cases for another $1 billion, FDAnews Drug Daily Bulletin, 1, September 28, 2004.
5. Shoemaker, C.J., A call to arms: the history of vaccines, Available at www.attorneyaccess.net
6. CDC, Update: vaccine side effects, adverse reactions, contraindications, and precautions reommendations of the advisory committee on immunization practices (ACIP), *MMWR*, 45(RR12;001), September 6, 1996.
7. *Reyes v. Wyeth Laboratories*, 498 F.2d 1264 (5th Cir. 07/31/1974).
8. Carl Vinson Institute of Government, The University of Georgia chronology of the 1918 Spanish Influenza Epidemic in Georgia, 1999. Available at www.cviog.uga.edu
10. Neustadt, R.E. and Fineberg, H.V., *The Swine Flu Affair: Decision-Making on a Slippery Disease, Department of Health, Education, and Welfare*, Washington, DC, 1978.
11. *The National Childhood Vaccine Injury Act of 1986*, Section 2125, Public Health Service Act as codified at 42 U.S.C. 300aa (Suppl.), 1987.
12. *Vaccine Injured Children's Compensation Act of 2001*, HR 1287, March 29, 2001.
13. *Bill Summary and Legislative Action: Vaccine Injured Children's Compensation Act of 2001*, HR 1287, March 29, 2001, last major action April 16, 2001, Available at http://thomas.loc.gov
14. *Homeland Security Act of 2002*, Pub. L. 107–296, January 24, 2003 (codified at 42 *USCS* § 233).
15. *Smallpox Emergency Personnel Protection Act of 2003*, Pub. L. 108-20, April 30, 2003 (retroactive effective date January 24, 2003).
16. Mair, J.S. and Mair, M., Vaccine liability in the era of bioterrorism, *Biosecurity Bioterror. Biodefens Strat. Pract. Sci.*, 1, 169, 2003.
17. CDER, *CDER Report to the Nation: 2003*, U.S. Government Printing Office, Washington, DC, 2003.
18. Pear, R., In a shift Bush moves to block medical suits, *New York Times*, S1, P1, C1, July 25, 2004.
19. Savitsky, M.E., Recognizing hospital adverse drug reactions, *J. Pharm. Pract*., 2, 203–208, 1989.
20. Physician Insurers Association of America, Medication Error Study, PIAA, Washington, DC, June 1993.
21. Jick, H., Adverse drug reactions: the magnitude of the problem, *J. Allergy Clin. Immunol.*, 74, 555–557, 1984.
22. Stetler, C.J., Drug-induced illness (letter), *JAMA*, 229, 1043–1044, 1974.
23. Miller, R.R., Interpretation of studies on adverse drug reactions, *Am. J. Hosp. Pharm.*, 34, 753–754, 1977.
24. Physician Insurers Association of America, Statistical data from the PIAA cumulative data sharing report, January 1, 1985–December 31, 1999, as reported in *Reporter*, Texas Medical Liability Trust, 1–8, September/October 1999.
25. Curtis, L.H., Ostbye, T., Sendersky, V., Hutchison, S., Dans, P.E., Wright, A., Woosley, R.L., and Schulman, K.A., Inappropriate prescribing for elderly Americans in a large outpatient population, *Arch. Intern. Med.*, 164, 1621–1625, 2004.

26. Lazarou, J., Pomeranz, B.H., and Corey, P.N., Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies, *JAMA*, 279, 1200–1205, 1998.
27. Lasser, K.E., Allen, P.D., Woolhandler, S.J., Himmelstein, D.U., Wolfe, S.M., and Bor, D.H., Timing of new black box warnings and withdrawals for prescription medications, *JAMA*, 287, 2215–2220, 2002.
28. Bakke, O.M., Manocchia, M., de Abajo, F., Kaitin, K.I., and Lasagna, L., Drug safety discontinuations in the United Kingdom, the United States, and Spain from 1974 through 1993: a regulatory perspective, *Clin. Pharmacol. Ther.,* 58, 108–117, 1995.
29. Brewer, T. and Colditz, G., Postmarketing surveillance and adverse drug reactions: current perspectives and future needs, *JAMA*, 281, 824–829, 1999.
30. Thase, M.E., How should efficacy be evaluated in randomized clinical trials of treatments for depression?, *J. Clin. Psychiatry,* 60(Suppl. 4), 23–31, 1999.
31. Lurie, P. and Sasich, L.D., Safety of FDA-approved drugs, *JAMA,* 282, 2297–2298, 1999.
32. Willman, D., How a new policy led to seven deadly drugs, *Los Angeles, Times,* Section A1, December 20, 2000.
33. Sasich, L., Lurie, P., and Wolfe, S.M., The drug industry's performance in finishing postmarketing research (phase IV) studies, letter and report from *Public Citizen, Health Research Group Publication #1520*, Washington, DC, April 13, 2000.
34. Peay, M.Y. and Peay, E.R., The role of commercial sources in the adoption of a new drug, *Soc. Sci. Med.,* 26, 1183–1189, 1988.
35. Stross, J.K., Information sources and clinical decisions, *J. Gen. Intern. Med*., 2, 155–159, 1987.
36. Jones, M.I., Greenfield, S.M., and Bradley, C.P., Prescribing new drugs: qualitative study of influences on consultants and general practitioners, *BMJ*, 323, 378–381, 2001.
37. Basara, L.R., The impact of a direct-to-consumer prescription medication advertising campaign on new prescription volume, *Drug Inf. J.*, 30, 715–729, 1996.
38. Petersen, M., Pushing pills with piles of money: Merck and Pharmacia in arthritis drug battle, *New York Times,* C1, October 5, 2000.
39. Hurwitz, M.A. and Caves, R.E., Persuasion or information? promotion and the shares of brand name and generic pharmaceuticals, *J. Law. Econ.*, 31, 299–320, 1988.
40. Murphy, M.N., Smith, M.C., and Juergens, J.P., The synergic impact of promotion intensity and therapeutic novelty on market performance of prescription drug products, *J. Drug Issues*, 22, 305–316, 1992.
41. Wardell, W.M., Tsianco, M.C., Anavekar, S.N., and Davis, H.T., Postmarketing Surveillance of New Drugs II, Case Study, *J. Clin. Pharm.*, April 1979.
42. Trontel, A., CDER/CBER PDUFA 3 Risk Management Working Group, April 9, 2003.
43. Schiff, G., Monitoring and reporting adverse drug reaction, *Infusion,* 9, l985.
44. Starfield, B., Is U.S. health really the best in the world?, *JAMA*, 284, 483–485, 2000.
45. Wood, A.J., Thrombotic thrombocytopenic purpura and clopidogrel: a need for new approaches to drug safety, *N. Engl. J. Med.*, 342, 1824–1826, 2000.
46. Specific requirements on content and format of labeling for human prescription drugs: warnings Title 21 *Code of Federal Regulations*, Part 201.57(e), pp. 23, revised 2003.
47. Wolfe, S.M., Letter from Health Research Group to Margaret Heckler, Department of Health and Human Services, Washington, DC, December 28, 1983.
48. Fincham, J., Drug classification of Mississippi adrs. a statewide program to stimulate reporting of adverse drug reactions, *J. Pharm. Pract.*, 2, 239–244, 1989.
49. Connors, P.J., Medication error study: a summary, *Physician Insurers Association of America Medication Error Study*, PIAA, Washington, DC, June 1993.
50. Who is responsible, and shy, for the chaotic confusion over the polio inoculations? *Harper's Magazine,* August 1995.
51. Bellanti, J.A., Ed., *Immunology III,* WB Saunders Co., Philadelphia, PA, 1985.
52. Jonville-Bera, A.P., Autret-Leca, E., Barbeillon, F., and Paris-Llado, J., Sudden unexpected death in infants under three months of age and vaccination status: a case-control study, *Br. J. Clin. Pharmacol.*, 51, 271–276, 2001.
53. Coulter, H.L. and Fisher, B.L., *DPT: A Shot in the Dark: Why the P in DPT May Be Hazardous to Your Health*, Avery Publishing, Vonore, TN, 1985.

54. Voelkel, N.F., Clarke, W.R., and Hegenbottam, T., Obesity dexfenfluramine, and pulmonary hypertension: a lesson learned? *Am. J. Resp. Crit. Med.*, 155, 786–788, 1997.
55. Havron, M.D., Diet drugs and pulmonary hypertension, *J. Fam. Pract.*, 43, 434–435, 1996
56. Cleare, A.J., Phentermine, psychosis, and family history, *J. Clin. Psychopharmacol.*, 16, 470–471, 1996.
57. Kokkinos, J. and Levine, S.R., Possible association of ischemic stroke with phentermine, *Stroke*, 24, 310–313, 1993.
58. Bostwick, J.M. and Brown, T.M., A toxic reaction from combining fluoxetine and phentermine, *J. Clin. Psychopharmacol.*, 16, 189–190, 1996.
59. Dukes, M.N.G., Ed., *Side Effects of Drugs Annual 8: A Worldwide Yearly Survey of New Data and Trends,* Elsevier, New York, l984.
60. Offerhaus, L., Nonsteroidal anti-inflammatory drugs in the treatment of rheumatic disorders: pitfalls and problems with the present plethora, *Neth. J. Med.*, 26, 173, 1983.
61. Stern, R.S. and Bigby, M., An expanded profile of cutaneous reactions to nonsteroidal anti-inflammatory drugs: reports to a specialty-based system for spontaneous reporting of adverse reactions to drugs, *JAMA*, 252, 1433–1437, 1984.
62. Goudie, B.M., Birnie, G.F., Watkinson, G., MacSween, R.N., Kissen, L.H., and Cunningham, N.E., Jaundice associated with the use of benoxaprofen letter, *Lancet*, 1, 959, 1982.
63. Taggart, H., Fatal cholestatic jaundice in elderly patients taking benoxaprofen, *Br. Med. J.*, 284, 1372, 1982.
64. Committee on Safety of Medicines, Suspension of products licenses for benoxaprofen, *Lancet*, 2, 396, 1982.
65. O'Donnell, J., Adverse drug reactions: extent of the problem and causality assessment, in O'Donnell, J.T., Ed., *Drug Injury: Liability, Analysis, and Prevention*, L&J Publishers, Tucson, AZ, 2005, pp. 81–97.
66. McEvoy, G.K., Ed., Drug Information 2004, *American Hospital Formulary System*, American Society of Hospital Pharmacists, Bethesda, MD, 2004, pp. 2896–2923.
67. Willman, D., Rezulin: Fast-Track Approval and a slow Withdrawal, *Los Angeles, Times,* Section A1, December 20, 2000. Fraser, C.G., Adrenal atrophy and irreversible shock associated with cortisone therapy, *JAMA*, 149, 1542–1543, 1952.
68. CNN, Diabetes drug Rezulin taken off market, March 23, 2000, Available at www.cnn.com/2000/HEALTH/03/22/diabetes.drug.01
69. Willman, D., Case study, Richard C. Eastman: A federal researcher who defended a client's lethal drug, *Los Angeles, Times,* Section A1, December 7, 2003.
70. WebMd, Medical News Archive, Drug alert: Rezulin removed from the market, March 21, 2000.
71. Willman, D., Drug tied to deaths is pulled; Safety: Baycol is a cholesterol-lowering medication that may lead to kidney failure, *Los Angeles, Times,* Section A1, August 9, 2001.
72. Jury Finds Rezulin diabetes drug not liable for deaths, Defeat Diabetes Foundation, Available at www.defeatdiabetes.org/articles/drugs040603.htm, posted June 3, 2004.
73. Lehmann, J., Statins and Baycol: questionable cholesterol control, in O'Donnell, J.T., Ed., *Drug Injury: Liability, Analysis, and Prevention*, 2005, pp. 469–472.
74. *Haltom v. Bayer Corp.*, No. 02-60165-00-0-4 (Nueces Co., TX).
75. Kauffman, J.M., *Malignant Medical Myths*. Infinity Publishing.com, West Conshohocken, PA, 2006, pp. 90, 92.
76. *Horn v. Thoratec Corp.*, No. 02-4597 (3d Cir. May 11, 2004).
77. Klein, A., Taylor, M.A., and Nelson, C.E., Another circuit bars tort suits under the MDA, Duane Morris, LLP, New York, NY, August 13, 2004.
78. *Buckman Co. v. Plaintiffs' Legal Comm.* (98-1768) 531 US 341 (2001) 159 F.3d 817, reversed, Argued December 4, 2000–Decided February 21, 2001.
79. *McCain-Kennedy-Edwards Patients' Bill of Rights* of 2001, S-283.
80. Johnson, D.R., Canadian imports will crimp drug development, *Bangor Daily News*, May 14, 2004.
81. Transcript of Congressman Hinchey's Remarks on the House Floor, July 13, 2004, Available at www.house.gov/hinchey/newsroom/071304_transcript.shtml

29

Problems in the Nondrug Marketplace

Stephen Barrett

CONTENTS

29.1 Overview

For purposes of discussion, drug and dietary supplement products can be divided into six categories: (1) prescription drugs, (2) standard over-the-counter (OTC) products, (3) vitamins and minerals, (4) other nutritional supplements, (5) herbs, and (6) homeopathic products. Prescription and standard OTC drugs are closely regulated and, almost always, live up to their advertised claims. Supplements and herbs are considerably less regulated, and many involve unsubstantiated claims, insufficiently tested ingredients, and poor quality control. Homeopathic products are minimally regulated. They are presumed safe, but there is no logical reason to use them. Table 29.1 compares the five product categories.

29.2 Prescription Drugs

Prescription drugs must have Food and Drug Administration (FDA) approval as safe and effective for their intended purposes and are tightly regulated. To gain approval, they must undergo laboratory tests and well-designed clinical trials that demonstrate safety

TABLE 29.1

Regulatory Status of Drugs and Dietary Supplements

	Prescription Drugs	OTC Drugs	Vitamins and Minerals	Other Nutritional Supplements	Herbal Products	OTC Homeopathic Products
Permissible claims	Must be generally recognized by experts as safe and effective for their intended purposes	Must be generally recognized by experts as safe and effective for their intended purposes	"Structure–function" claims permitted, but drug claims require premarket approval	"Structure–function" claims permitted, but drug claims require premarket approval	"Structure–function" claims permitted, but drug claims require premarket approval	Can be marketed for treatment of self-limiting ailments readily diagnosable by consumers
Labeling	Must provide ingredient amounts and adequate directions for use, including warnings	Must provide ingredient amounts and adequate directions for use, including warnings	Must list serving size, ingredients, and quantities of vitamin/mineral ingredients. No warnings are required	Must list serving size, ingredients, and quantities of vitamin/mineral ingredients. No warnings are required	Must list ingredients but not quantities. No warnings are required	Must identify ingredients and provide directions for use; products are assumed to be nontoxic
FDA approval	New Drug Application (NDA) must be approved	Must conform to OTC monographs	Not required for truthful structure–function claims	Not required for truthful structure–function claims	Not required for truthful structure–function claims	Not required for use of ingredients listed in the U.S. Homeopathic Pharmacopeia
Advertising regulation	FDA (stringent regulation; violative claims are uncommon)	FTC (violative claims are rare)	FDA/FTC	FDA/FTC (violations greatly exceed regulatory resources)	FDA/FTC (violations greatly exceed regulatory resources)	FDA/FTC (minimal regulation)
Adverse event reporting	Required	Required	Not required	Not required	Not required	Not required
GMP required	Yes	Yes	Yes, but no standards have been published	Yes, but no standards have been published	Yes, but no standards have been published	Yes
Product quality	Generally good	Generally good	Generally good	Inconsistent	Poor; most products have not been standardized	No data exist
Recommended legislative and regulatory reforms	None needed	None needed	FDA should be authorized to limit dosages	Should be regulated as least as stringently as OTC drugs	Should be regulated as least as stringently as OTC drugs	Public warning or ban should be implemented

and effectiveness. The testing and approval of a new drug is very expensive and takes many years.

Before a new drug can be approved and marketed, the manufacturer generally must take the following steps:

- The drug must be subjected to laboratory and animal tests, which must indicate that it can be safely tested in humans.
- Before the drug is given to people, the sponsor must submit a "Notice of Claimed Investigational Exemption for a New Drug," which is commonly referred to as an "investigation new drug" (IND). The application must describe the composition of the substance, the results of the animal studies, the design (protocol) of the proposed clinical trial, the measures that will protect the experimental subjects, and the training and experience of the investigators.
- Approved clinical investigations follow three phases. In Phase 1, about 100 to 200 people are exposed to the drug to determine the tolerance, absorption, excretion, half-life, and other pharmacologic reactions; the preferred route of administration; and the safe dosage. In Phase 2, initial trials are conducted on 500 to 1000 patients to assess the treatment or prevention of the specific disease. Additional animal studies to indicate safety may be conducted concurrently. If these preliminary studies demonstrate sufficient promise, Phase 3 clinical trials are performed with several thousand patients.
- The manufacturer submits a new drug application (NDA) that includes the experimental data, a sample package insert, and the proposed label.
- The FDA approves the NDA, asks for further evidence, or rejects the application.

In Phase 3 clinical trials, the experimental group of people who receive the treatment being tested are compared with a control group of people who receive a different treatment or no treatment. For example, the experimental group may receive tablets with active ingredients, whereas the control group receives another treatment, an inert substance (placebo), or no treatment. In "double-blind" trials, neither the experimenters nor the patients know who gets what. Controlled studies are necessary to distinguish between drug effects, placebo effects, and the natural course of the ailment. A placebo effect is a beneficial response to a substance, device, or procedure that cannot be accounted for on the basis of pharmacologic or other direct, physical action. Feeling better when the physician walks into the room is a common example.

After marketing begins, manufacturers must report adverse effects. If a drug produces unexpected side effects or is found to be less effective than expected, approval can be withdrawn, which happens from time to time.

29.3 Over-the-Counter Drugs

Over-the-counter drugs must be generally regarded by experts as safe and effective and suitable for self-use by consumers. In 1972, the FDA began an extensive evaluation to ensure the effectiveness and safety of OTC drugs. Rather than attempting to evaluate hundreds of thousands of individual products, it divided the ingredients into categories and appointed expert advisory panels to evaluate the categories. The first phase of the review spanned about 10 years and covered 722 ingredients. After advisory panels concluded that

only one third of these ingredients were safe and effective, manufacturers reformulated many products by removing unsafe and ineffective ingredients or by adding others. In addition, some prescription drug ingredients were approved for use in OTC drugs, generally in lower dosages. Today, except for homeopathic products, the vast majority of OTC drugs are backed by scientific evidence that they are safe and effective for their intended purposes.

29.4 Vitamins, Minerals, and Other Nutritional Supplements

The *Federal Food, Drug, and Cosmetic Act* defines a "drug" as any article (except devices) "intended for use in the diagnosis, cure, mitigation, treatment, or prevention of disease" and "articles (other than food) intended to affect the structure or function of the body." Drugs that are not "generally recognized as safe and effective by experts" are "new drugs" that cannot be legally marketed without FDA approval. The pertinent laws and regulations permit the FDA to stop the marketing of "dietary supplements" with unsubstantiated drug claims on their labels.

To evade the law's intent, sellers use many channels to ensure that the public learns of "medicinal" uses that are not stated on product labels. This is done mainly by promoting the ingredients of the products in books, magazines, newsletters, booklets, lectures, radio and television broadcasts, oral claims by retailers, and communication through the Internet. Although these communications assert that various substances can function like drugs, the FDA lacks jurisdiction unless the claims are made directly in the labeling or advertising of the products.

Passage of the *Dietary Supplement Health and Education Act of 1994* (DSHEA) worsened this situation by increasing the amount of misinformation that can be *directly* transmitted to prospective customers. It also expanded the types of products that could be marketed as "supplements." The most logical definition of dietary supplement would be something that supplies one or more essential nutrients missing from the diet. Dietary Supplement Health and Education Act went far beyond this to include vitamins, minerals, herbs or other botanicals, amino acids, other dietary substances to supplement the diet by increasing dietary intake, and any concentrate, metabolite, constituent, extract, or combination of any such ingredients. Although many such products (particularly herbs) are marketed for their alleged preventive or therapeutic effects, DSHEA has made it difficult or impossible for the FDA to regulate them as drugs. Since its passage, even hormones, such as dehydroepiandrosterone (DHEA), human growth hormone (HGH), and melatonin, are hawked as supplements. The law does not require dietary supplements to be proven effective before marketing.

Dietary Supplement Health and Education Act also prohibits the FDA from banning dubious supplement ingredients as "unapproved food additives." The FDA considered this strategy more efficient than taking action against individual manufacturers who made illegal drug claims. Because of the passage of DSHEA, the only way to banish an ingredient is to prove it unsafe. Ingredients that are useless but harmless are protected.

The marketplace would be safer if the FDA could regulate the upper dosage limits of vitamins and minerals and could hold other nutritional supplements to at least the same standards as those for OTC drugs. But DSHEA prevents the FDA from doing this unless a product poses an imminent hazard.

As manufacturers are not required to submit safety information before marketing dietary supplements, the FDA must rely on adverse event reports, product sampling,

information in the scientific literature, and other sources for evidence of danger. Because the FDA is unable to monitor and regulate thousands of individual products, the public is virtually unprotected against supplements and herbs that are unsafe.

The agency is permitted to restrict substances that pose a "significant and unreasonable risk" under the conditions of use on the label or as commonly consumed. The FDA has the burden of proof and cannot act until substantial harm occurs to consumers, but manufacturers are not required to report adverse events. Even when harm is obvious, instituting a ban can take years. For example, even though thousands of consumers appear to have been affected by products that contain ephedra, it took the FDA more than 5 years to go through the procedures needed to ban its use.

The Federal Trade Commission (FTC) has jurisdiction over the advertising of nonprescription products that are marketed in interstate commerce. Although the FTC has a very powerful law, the agency can handle only a small percentage of the violations it detects. The Postal Service (USPS) has overlapping jurisdiction over products sold by mail. The USPS also has a strong law but has not pursued falsely advertised health products for more than 10 years.

29.5 "Nutritional Support" Statements

Dietary Supplement Health and Education Act allows dietary supplements to bear "statements of support" that (1) claim a benefit related to classical nutrient-deficiency disease; (2) describe how ingredients affect the structure or function of the human body; (3) characterize the documented mechanism by which the ingredients act to maintain the body's structure or function; and (4) describe general well-being from consumption of the ingredients. To be legal under DSHEA, a nutritional support statement must not be a drug claim. In other words, it should not suggest that the product or ingredient is intended for the prevention or treatment of disease. However, the marketplace has been flooded by statements related to organs (such as "supports the eyes" or "supports the cardiovascular system"), which deliver an equivalent message.

Actually, few statements about the biochemical or physiological properties of nutrients have practical value for consumers. By definition, every essential nutrient is important to proper body function. Simple statements about nutrient function are more likely to be misleading than helpful. A statement such as "vitamin A is essential to good eye function" could suggest that (1) people need to take special steps to be sure they get enough; (2) extra vitamin A may enhance eyesight; and (3) common eye problems may be caused by vitamin A deficiency or remedied by taking supplements. To be completely truthful, a "nutritional support" statement about vitamin A would have to counter all three misconceptions and indicate that people eating sensibly do not need to worry about whether their vitamin A intake is adequate. In other words, truthful statements about nutrient supplements would have to indicate who does *not* need them. No vitamin manufacturer has ever done this or ever will. Because herbs are not nutrients, the concept of nutritional support statements for herbs is absurd. Yet, thousands of herbal products are marketed with claims that they support various body functions.

Under DSHEA, manufacturers who make statements of nutritional support must have substantiation that such statements are truthful and not misleading. The law also requires that the Secretary of Health and Human Services (HHS) be notified no later than 30 days after the first marketing of a supplement for which the statement is made. The law does not define substantiation.

Historically, the FDA has considered literature used directly in connection with the sale of a product to be "labeling" for the product. Dietary Supplement Health and Education Act exempts publications from labeling if they (1) are not false or misleading, (2) do not promote a particular manufacturer or brand, (3) present a balanced view of pertinent scientific information, and (4) are physically separated from the items discussed. However, because substantial percentages of dietary supplements are useless, irrationally formulated, or overpriced, the supplement industry has little reason to provide literature that is not misleading. In addition, the FDA does not have the resources to police the huge numbers of "support" statements to ensure that they are appropriately "balanced."

The passage of DSHEA was spearheaded by an avalanche of communication to Congress from people who expressed fears that without special protection consumers would lose their freedom to buy vitamins. These fears were unfounded, but many congressional representatives concluded that the FDA should be restrained.[1]

Ironically, by the time DSHEA was enacted, the OTC drug review had rid the marketplace of nearly all the standard OTC products that had been marketed with unsubstantiated claims. In contrast, few dietary supplements promoted with disease-related claims can live up to such claims. A few years ago, a review of 964 herbs and dietary supplement products listed in the Natural Medicines Comprehensive Database found that only 15% had been proven safe and only 11% had been proven effective for the indications for which they were advocated.[2]

A Roper survey of 1480 persons age 50 or older, conducted in 2001, found that about 75% of the respondents wanted the government to review safety data and approve dietary supplements before sale and to verify all health-related claims before they can be included in advertisements and on product labels.[3] This is just contrary to what DSHEA permits, but there is little hope that Congress will modify its provisions.

29.6 Herbal Problems

Many herbs contain hundreds or even thousands of chemicals that have not been completely cataloged. Some of these chemicals may turn out to be useful as therapeutic agents, but others could prove harmful. No legal standards exist for the processing, harvesting, or packaging of medicinal herbs, and the FDA does not require herbal products to adhere to any standards of identity or dosage. Many investigations have found that the amounts of the ingredients listed on herbal product labels are inaccurate. Moreover, many herbal practitioners are nonphysicians who are not qualified to make appropriate diagnoses or to determine how herbs compare with proven drugs. Thus, even if a substance is potentially useful, consumers cannot be certain what is in the finished products or how to use them.

Herbs in their natural state can vary greatly from batch to batch and often contain chemicals that cause side effects but provide no benefit. Surveys conducted in the United States have found that the ingredients and doses of various products vary considerably from brand to brand and even between lots of the same product.[4] For example, researchers at the University of Arkansas tested 20 supplement products containing ephedra (ma huang) and found that half the products exhibited discrepancies of 20% or more between the label claim and the actual content, and one product contained no ephedra alkaloids. Ephedra products are marketed as "energy boosters" or "thermogenic" diet aids, even though no published clinical trials substantiate that they are safe or effective for these purposes. The researchers also noted that hundreds of such products are marketed and that their number

exceeds that of conventional prescription and nonprescription ephedra products, which are FDA approved as decongestants.[5] Studies have also found that the recommended dosage can vary greatly from product to product.

To make a rational decision about a herbal product, it would be necessary to know what it contains, whether it is safe, how it interacts with drugs or other herbs, and whether it has been demonstrated to be as good or better than pharmaceutical products available for the same purpose. For most herbal ingredients this information is incomplete or unavailable.

The National Council Against Health Fraud (Peabody, MA, U.S.A.) believes that the FDA could improve this situation by establishing a special category of OTC drugs called "traditional herbal remedies." The council's proposal would permit these products to be marketed with less-than-standard proof of effectiveness provided that (1) reasonable evidence exists that they are safe and effective; (2) labels identify the name and quantity of each active ingredient; (3) indications are restricted to nonserious, self-limiting conditions; (4) labels contain adequate directions for use, including a warning about inappropriate self-treatment; and (5) adverse reactions are reported.[6] These standards are similar to the current OTC standards. However, even if the products were improved, only a few herbs have been studied enough to conclude that they might be effective, and even fewer have been studied enough to determine whether their use is practical.

Garlic provides an example of research-based support that does not necessarily make its use practical. Although some garlic preparations have been demonstrated to lower cholesterol, prescription drugs are more potent for this purpose, and garlic's anticoagulant properties could pose a significant risk. No data are available to indicate the risk of combining garlic with other widely used products (vitamin E, ginkgo, fish oil, or aspirin, for instance) that can interfere with blood clotting.

Adulteration, which can be accidental or deliberate, is another problem. Many herbal products have been found to contain prescription or OTC drugs and dangerous heavy metals. In 1998, for example, the California Department of Health reported that 32% of Asian herbal medicines sold in that state contained undeclared pharmaceuticals or heavy metals.[7] A subsequent study of more than 500 Chinese herbal medicines found that about 10% of them contained undeclared drugs or toxic levels of metals.[8] The FDA and other investigators have also detected sildenafil, colchicine, adrenal steroids, alprazolam, and other prescription drug ingredients in products claimed to contain only natural ingredients.

Although drugs are subject to rigorous quality control to ensure the nature, potency, and safety of their ingredients, dietary supplements and herbs are not.[9] The FDA has proposed adherence to good manufacturing practices (GMPs) but has not taken the necessary steps to require this.

ConsumerLab (White Plains, NY, U.S.A.), which has tested hundreds of products since 1999, reported that about 15% of the vitamins and minerals, 23% of the other supplements, and 38% of the herbals failed their evaluations. The most common reason for the failure was too little or none of the main ingredient. The other problems included too much active ingredient, the wrong ingredient, dangerous or illegal ingredients, contamination with heavy metals, pesticides or pathogens, "spiking" with unexpected ingredients, poor disintegration (which affects absorption), and misleading product information.[10]

The entry of drug companies into the herbal marketplace may result in dosage standardization for some products, and public and professional interests in herbs are likely to stimulate more research. However, with safe and effective medicines available, treatment with herbs rarely makes sense, and many of the conditions for which herbs are recommended are not suitable for self-treatment.

29.7 Homeopathic Products

Homeopathy dates back to the late 1700s, when a German physician named Samuel Hahnemann began formulating its basic principles. Proponents postulate that (1) diseases represent a disturbance in the body's ability to heal itself, (2) only a small stimulus is needed to begin the healing process, (3) substances that produce symptoms in healthy people can cure similar symptoms in sick people ("law of similars"), and (4) remedies become more potent with greater dilution. The word "homeopathy" is derived from the Greek words *homoios* (similar) and *pathos* (suffering or disease).

Homeopathic products are derived from minerals, botanical substances, zoological substances, microorganisms, and several other sources. Soluble sources are serially diluted with distilled water or alcohol. Insoluble sources are pulverized and mixed with milk sugar. One part of the diluted mixture is diluted again, and the process is repeated to reach the desired concentration. Serial dilutions of 1:10 are designated by the roman numeral "X" (e.g., 1X = 1/10, 3X = 1/1,000, and 6X = 1/1,000,000). Serial dilutions of 1:100 are expressed with the roman numeral "C" (1C = 1/100, 3C = 1/1,000,000, and so on.) Most products range from 6X to 30C, but some carry designations as high as 1000X.

Dilution beyond Avogadro's number (6.022×10^{23}) results in the loss of the original substance altogether. This limit corresponds to 12C or 24X potencies (one part in 10^{24}). But proponents claim that vigorous shaking ("succussion") or pulverizing between dilutions would leave behind a spirit-like essence that cures by reviving the body's "vital force." Proponents also postulate that the solute retains a "memory" of the substance. If this were true, every substance encountered by a molecule of water, alcohol, or milk sugar might imprint an "essence" that could exert powerful and unpredictable medicinal effects.

Many proponents allege that homeopathic products resemble vaccines because both provide a small stimulus that triggers an immune response. This comparison is not valid. The amount of active ingredients in vaccines is much greater and is measurable. Immunizations produce measurable antibodies, but ultradilute homeopathic products have no measurable active ingredients and produce no measurable response. In addition, vaccines are used preventively, not for curing symptoms.

Homeopathy was given legal status by the 1938 *Federal Food, Drug, and Cosmetic Act*, whose chief sponsor, Senator Royal Copeland, was also a homeopathic physician.[11] One provision of this law recognizes all substances included in the *Homœopathic Pharmacopœia of the United States* as drugs.[12] The current (ninth) edition of this book provides protocols for preparing more than 1200 substances for homeopathic use but says nothing about how those substances should be used. Manufacturers and practitioners decide that.

The basis for formulating homeopathic products is not clinical testing but homeopathic "provings," during which substances are administered to healthy people who record their thoughts and physical sensations over various periods of time. The reported symptoms have been compiled into lengthy reference books, called *Materia Medica*, which proponents regard as gospel and use to match the patient's symptoms to "corresponding" products. However, most of the provings were done between 100 and 200 years ago, when medical science was in its infancy. Very little was known about the nature of health and disease or about how to conduct experiments that separate cause and effect from coincidence.

The fact that a symptom occurs after taking a substance can have several explanations. During a typical day, most people experience occasional unpleasant thoughts and bodily sensations. To determine whether a substance actually causes a symptom, it would be necessary that people who receive the substance be compared with those who receive a

control substance. To guard against bias, neither the experimenters nor the test subjects should know who received what.

The provings used to compile *Materia Medica* were not conducted in the aforementioned way. There were wide variations in the amount of substances administered, the timing of the administrations, the way in which data were recorded, and the length of the studies; in addition there were no controls.[13] Thus, it is impossible to know whether the reported symptoms were actually related to the administration of the test substances. In addition, many symptoms may have resulted from the suggestibility of the test subjects. In most cases, participants have known which substances were administered. Potencies have ranged from the undiluted original substance to dilutions as high as 200C. Provings also constitute the basis for inclusion in the *Homœopathic Pharmacopœia.*

A Dictionary of Practical Materia Medica,[14] a widely used three-volume set authored by John Henry Clarke, illustrates the foolishness involved in provings. The book contains about 2500 pages that describe the symptoms that supposedly were reported following administration of about 1200 substances. Most descriptive pages contain more than 100 claims, which means that the total number of symptoms exceeds 200,000. The book does not indicate when or how the original provings were done or who reported most of the specific findings. Thus, it would be impossible to examine whether the studies were properly done, who did them, and whether the findings were accurately reported.

Many of the listed symptoms are odd. *Lac felinum* includes "cannot bear the smell of clams, of which she is naturally fond;" *Latrodectus mactans* "screams fearfully, exclaiming that she would lose her breath and die;" *Magnesia sulphurica* "stupidity;" *Oleum animale* "singing, tinkling, and buzzing in ears;" *Natrum carbonicum* "hurries out of bed in the morning." Some listings include symptoms that occur predominantly on one side of the body, such as "sickening sensation in left testicle." All are supposedly useful in determining whether the patient might "fit" a particular remedy.

Even if the proving reports were consistent, there is neither a logical reason why substances that could produce symptoms should cure such symptoms nor evidence from appropriately designed studies that the "law of similars" actually operates. The real way to test whether something works is to test whether it helps sick people. This requires clinical trials in which people who receive the test substance are compared with people who do not. No homeopathic product has ever been proven effective; and a vast majority of products have never even been clinically tested.

"Classical" homeopaths purport to tailor their prescriptions to the individual patient on the basis of a detailed history. Some practitioners maintain that certain people have special affinity to a particular remedy (called their "constitutional remedy") to which they will respond for a variety of ailments. However, most practitioners prescribe products targeted at symptoms rather than at the individual's "constitution." Some practitioners select remedies with the help of a galvanometer purported to measure "electromagnetic energy imbalances" responsible for disease.[15] Some proponents rationalize that homeopathic remedies are not testable like standard drugs because treatment must be individualized. However, most OTC homeopathic products are mixtures marketed directly to consumers.

In 1995, *Prescrire International*, a French journal that evaluates pharmaceutical products, published a review that concluded: "As homeopathic treatments are generally used in conditions with variable outcome or showing spontaneous recovery (hence their placebo responsiveness), these treatments are widely considered to have an effect in some patients. However, despite the large number of comparative trials carried out to date there is no evidence that homeopathy is any more effective than placebo therapy given in identical conditions."[16]

In 1996, a report by the Homeopathic Medicine Research Group (HMRG, formed by the European Commission [Brussels, Belgium]) evaluated 184 published and unpublished reports of controlled trials of homeopathic treatment. The panelists concluded that (1) only

17 trials were designed and reported well enough to be worth considering; (2) in some of these trials, homeopathic approaches may have exerted a greater effect than a placebo or no treatment; and (3) the number of participants in these trials was too small to draw any conclusions about the effectiveness of homeopathic treatment for any specific conditions.[17]

Other reviewers, while asserting that the clinical effects reported in 89 randomized, controlled trials were greater than expected from placebo, acknowledged that homeopathy has not been proven effective for any single, clinical condition.[18] But even if results could be consistently produced, most of the study of a single remedy for a single disease could prove is that the remedy is effective against that disease. It would not validate homeopathy's basic theories or prove that homeopathic treatment is useful for other diseases.

In 1999, *The Medical Letter on Drugs and Therapeutics* concluded that "the chemical content of homeopathic products is often undefined, and some are so diluted that they are unlikely to contain any of the original material. These products have not been proven effective for any clinical condition. There is no good reason to use them."[19]

Although the FDA could insist that homeopathic products be proven effective by scientific testing to remain marketable, it has not chosen to do so. Its regulations merely state that homeopathic drugs cannot be offered without prescription for such serious conditions as cancer, AIDS, or any other requiring diagnosis and treatment by a licensed practitioner. Nonprescription homeopathics may be sold only for self-limiting conditions recognizable by consumers.[20] No homeopathic product has FDA approval as a result of having an NDA approved. Products offered for conditions not amenable to OTC use can be marketed as prescription products, the labels of which generally make no claims.

If the FDA required homeopathic remedies to be proven effective in order to remain marketable — the standard it applies to other categories of drugs — homeopathy would face extinction in the United States.[21] However, there is no indication that the agency is considering this. FDA officials regard homeopathy as relatively benign (compared, for example, to unsubstantiated products marketed for cancer and AIDS) and believe that other problems should get enforcement priority. If the FDA attacks homeopathy too vigorously, its proponents might even persuade a lobby-susceptible Congress to rescue them.

In 1994, 42 prominent critics of quackery and pseudoscience asked the agency to curb the sale of homeopathic products.[22] The petition urged the FDA to initiate a rule-making procedure to require that all OTC homeopathic drugs meet the same standards of safety and effectiveness as nonhomeopathic OTC drugs. It also asked for a public warning that although the FDA has permitted homeopathic remedies to be sold, it does not recognize them as effective. The FDA did not officially respond to the petition. However, in 1998, at a symposium sponsored by *Good Housekeeping* magazine, former FDA Commissioner David A. Kessler, acknowledged that homeopathic remedies do not work but that he did not attempt to ban them because he felt that Congress would not support a ban.[23]

29.8 Recommended Reforms

Government agencies have never had sufficient resources to cope with the enormous amount of deception in the marketing of herbs, dietary supplements, and homeopathic products. Dietary Supplement Health and Education Act made the problem worse. Without an adequate law, the FDA cannot curb the deceptive marketing of what DSHEA calls dietary supplements. Regulators have the power to ban homeopathic remedies but have shown no interest in doing so.

I believe, that in matters of health, there should be no tolerance for wrongdoing. In line with that philosophy, the following measures are needed:

- Dietary Supplement Health and Education Act should be repealed.
- The FDA should be authorized to set safe upper limits for vitamin dosage.
- Other dietary supplements and herbal remedies should be held to regulatory standards that are at least as stringent as those for OTC drug products. At the very least, active ingredients should be required to meet standards for purity, content, safety, and some level of effectiveness before they can be marketed.
- Homeopathic products should be banned. If that is not politically feasible, the agency should issue a public warning that none of them has been proven effective.
- At present, marketers who make unsubstantiated health claims are not penalized when the FDA warns them to stop. The FDA should be permitted to generate civil penalties in connection with warning letters.

29.9 Recommended Information Sources

Several sources provide reliable information on dietary supplements, herbs, and homeopathic products. Some recommended information sources include the following:

- The Natural Medicines Comprehensive Database[24] contains more than 1000 monographs. It is available online or in print for $92 per year (or $132 for both versions). It provides information about more than 1000 herbal and dietary supplement ingredients.
- About Herbs,[25] a project of Memorial Sloan-Kettering Cancer Center (New York, New York), evaluates more than 300 herbs, dietary supplements, and "alternative" cancer treatments. Both "professional" and consumer versions are provided, but most of the professional information is readily understandable by laypeople. The articles are shorter than their counterparts in the Natural Medicines database, but they are researched and written quite well.
- ConsumerLab.com, which has moderately priced subscriptions for more detailed information, posts product evaluations that include the results of laboratory tests of ingredient levels.
- *Herbal Medicines: A Guide for Healthcare Professionals*[26] contains 148 detailed, herbal monographs.
- *The Professional's Handbook of Complementary & Alternative Medicines*[27] provides practical advice for about 300 herbs and other substances.

References

1. Barrett, S. and Herbert, V., *The Vitamin Pushers: How the Health Food Industry Is Selling America a Bill of Goods*, Prometheus Books, Amherst, NY, 1994.
2. Marty, A.T., Natural medicines comprehensive database, *JAMA*, 283, 2992, 2000.
3. Prevention Magazine staff, *Survey of Consumer Use of Dietary Supplements*, Rodale Press, Emmaus, PA, 1999.

4. Barrett, S., The herbal minefield, Quackwatch Web Site, June 22, 2000, Available at www.quackwatch.org/01QuackeryRelatedTopics/herbs.html
5. Gurley, B.J. et al., Content versus label claims in ephedra-containing dietary supplements, *Am. J. Health Syst. Pharm.*, 57, 963, 2000.
6. NCAHF position paper on over-the-counter herbal remedies, National Council Against Health Fraud Web Site, Peabody, MA, 1995, Available at www.ncahf.org/pp/herbal.html
7. Ko, R.J., Adulterants in Asian patent medicines, *N. Engl. J. Med.*, 339, 847, 1998.
8. Au, A.M. et al., Screening methods for drugs and heavy metals in Chinese patent medicines, *Bull. Environ. Contam. Toxicol.*, 65, 112, 2000.
9. Rhoades, R.A., Global corner: dietary supplement GMPs and the global market, *Update*, 6, 41–42, 2001.
10. Cooperman, T., Obermeyer, W., and Webb, D., *ConsumerLab.com's Guide to Buying Vitamins & Supplements: What's Really in the Bottle?,* ConsumerLab.com, White Plains, NY, 2003.
11. Sjunod, S.W., An alternative perspective: homeopathic drugs, Royal Copeland, and federal drug regulation, *Pharmacy in History*, 42, 13, 2000, Available at www.homeowatch.org/history/reghx.html
12. Homeopathic Pharmacopœia Revision Service, *Homeopathic Pharmacopœia Convention of the United States (HPCUS)*, Homœopathic Pharmacopœia of the United States, Southeastern, PA, 1988–1994.
13. Sherr, J.Y., *The Dynamics and Methodology of Provings*, Dynamis Books, West Malvern, England, 1994.
14. Clarke, J.H., *A Dictionary of Practical Materia Medica*, B. Jain Publishers, New Delhi, India, 1921 (reprinted 2000).
15. Barrett, S., "Electrodiagnostic" devices, Quackwatch Website, revised February 13, 2004, Available at www.quackwatch.org/01QuackeryRelatedTopics/electro.html
16. Prescrire International, *la revue Prescrire*, 15(155), 674, 1995. [English translation in *Prescrire International*, 5(21), 1996], Available at http://www.hutch.demon.co.uk/prescrire
17. Homeopathic Medicine Research Group, *Report to the European Commission directorate general XII: science, research and development,* European Commission, Brussels, Belgium, 1996.
18. Linde, K. et al., Are the clinical effects of homeopathy placebo effects? A meta-analysis of placebo-controlled trials, *Lancet*, 350, 834, 1997.
19. Homeopathic products, *Med. Lett. Drug. Ther.*, 41, 20–21, 1999.
20. Office of Regulatory Affairs, *Conditions Under Which Homeopathic Drugs May Be Marketed*, CPG 7132.15, Sec. 400.400, FDA, Rockville, MD, Revised March 1995.
21. Pinco, R.G., *Status of Homeopathy in the United States: Important Ominous Developments*, Memo to Williard Eldredge, President, American Association of Homeopathic Pharmacists, January 17, 1985, Available at www.quackwatch.org/01QuackeryRelatedTopics/homeopetition/pinco.html
22. Barrett, S. et al., Petition regarding homeopathic drugs, Docket # 94P-0316/CP 1, Filed August 29, 1994.
23. Kessler, D.A., Panel discussion on herbal dietary supplements, Consumer Safety Symposium on Dietary Supplements and Herbs, New York, March 3, 1998.
24. Therapeutic Research Faculty, Natural Medicines Comprehensive Database, Available at www.naturaldatabase.com
25. Memorial Sloan-Kettering Cancer Center, Information Resource: About Herbs, Botanicals, and Other Products, Available at www.mskcc.org/mskcc/html/11570.cfm
26. Barnes, J., Anderson, L.A., and Phillipson, J.D., *Herbal Medicines: A Guide for Healthcare Professionals*, Pharmaceutical Press, London, U.K., 2002.
27. Fetrow, C.W. and Avila, J.R., *The Professional's Handbook of Complementary & Alternative Medicines*, Springhouse Corporation, Springhouse, PA, 1999.

30

Patents and New Product Development in the Pharmaceutical and Biotechnology Industries

Henry Grabowski

CONTENTS

30.1 Introduction

This chapter*[1] examines the rationale for intellectual property protection in the development of new pharmaceutical products. Prior survey studies of R&D have found that patents play a more critical role in appropriating the benefits of innovation in pharmaceuticals compared to other high tech industries. This paper considers why this is so, based on an analysis of the economic characteristics of R&D costs and returns in the pharmaceutical and biotechnology industries. The final section examines recent policy developments and issues surrounding patent lifetime and generic competition in this industry.

Grilliches, in a 1992 survey paper, found that high social returns to research and development (R&D) are a major factor underlying the growth in per capita income and consumer welfare during the 20th century.[2] Many of the studies done by economists on this topic have found that the social returns to R&D are more than twice the private returns. A primary reason for this finding is that positive externalities are generally associated

* This chapter is based on a presentation to the Dallas Federal Reserve Bank in April 2002. A similar version of the paper was published by Georgetown Public Policy Review and by the Dallas Fed in *Science and Cents: The Economics of Biotechnology.*

with industrial innovations. As Scherer[3] stated in his leading graduate text in industrial organization, "Making the best use of resources at any time is important. But in the long run it is dynamic performance that counts" (p. 407).

The pharmaceutical and biotechnology industries, which are among the most research-intensive industries, have been the focus of several studies on cost-benefit and social return on R&D studies. Cutler and McClellan[4] surveyed a number of studies that investigate the impact of new drugs on increased longevity, worker productivity, and savings in other types of medical expenditures. They find significant aggregate net benefits to society from new drug introductions. Their analysis is consistent with cost–benefit analyses targeted toward specific medical conditions such as cardiovascular disease, depression and infectious disease. These studies have also found high incremental social benefits from new drug innovation (see, e.g., Triplett, 1999).[5]

Another general finding of the academic literature is that public policy actions can have a significant influence on the rate of innovation in particular industries. Among the key industry policies influencing the innovative process in pharmaceuticals are the public support of biomedical research, patents, Food and Drug Administration regulatory policy, and government reimbursement controls.[6] The focus of this paper is the role and impact of patents and intellectual property protection in the discovery and development of new pharmaceutical and biotechnical products.

The importance of patents to pharmaceutical innovation has been reported in several cross-industry studies. Levin et al.[7] and Cohen et al.[8] have undertaken surveys of the U.S. R&D managers in a large cross-section of industries, to identify the most important and necessary factors in appropriating the benefits from innovations. These factors included the competitive advantages of being first in the market, superior sales and service efforts, the secrecy and complexity of production and product technology, as well as patents. Both studies found that the pharmaceutical industry placed the highest importance on patents. But in contrast, many other research-intensive industries such as computer technologies, scientific instruments, and semiconductors, placed greater stress on factors like lead time and "learning by doing" efficiencies in production accruing to first movers. This reflects the fact that R&D investment periods and product life cycles are typically much shorter in these industries. Furthermore, their new products often involve complex systems of many components as opposed to the discrete nature of new biological and new chemical entities (NCEs). As a consequence, the costs of imitation relative to the costs of innovation are much higher in many other high-tech industries compared to pharmaceuticals.

The findings of these studies are in accordance with an earlier study performed by the British economists Taylor and Silberston.[9] Based on a survey of R&D managers in the United Kingdom, they estimated that pharmaceutical R&D expenditures would be reduced by 64% in the absence of patent protections. By contrast, the corresponding reduction was only 8% across all industries. Similar findings were reported by Edwin Mansfield, in a survey of the research directors of 100 U.S. corporations.† [10,11]

Some individuals have called for the abolishment of pharmaceutical parents on the grounds that they give rise to excessive profits and high prices on new medicines.[12] However, the suggestion that the government could replace the $27 billion R&D effort of

† In a follow-on study, Silberston categorized three groups of industries for when patents are essential, very important or less important on the basis of both survey responses and objective analyses (patent and R&D intensity). He concluded that, "The first category consists of one industry only, pharmaceuticals."[10] Edwin Mansfield surveyed the R&D director of 100 U.S. corporations on what fraction of the inventions they introduce, between 1981 and 1983 would not have been developed without patent protection. For pharmaceuticals, the value was 60%, while the average across all industries was 14%.[11]

the private pharmaceutical industry, and produce an equivalent stream of new pharmaceutical products is highly problematic on both economic and public policy grounds. Opponents of drug patents also ignore the fact that new products lead to social welfare gains for consumers, even when supplied by a single firm under a patent exclusivity grant. In addition, price competition occurs in the pharmaceuticals market under the current system when closely substitutable medicines in the same therapeutic family are introduced. Finally, there is intensive price competition when the originating brand's patents expire and generic entry occurs.

The following sections of this paper examine the economic characteristics of the R&D process in pharmaceuticals that make patents so critical. Sections 30.2 to 30.4 discuss the costs of innovation in pharmaceuticals and the effects on innovative and imitative competition of the 1984 Hatch–Waxman Act. Section 30.5 considers whether the biotech industry is different from the pharmaceutical industry in terms of R&D costs. Section 30.6 considers the distribution of returns on R&D in these industries. The final section presents conclusions and policy considerations.

30.2 R&D Costs for a New Drug Introduction

The explanation for why patents are more important to pharmaceutical firms in appropriating the benefits from innovation follows directly from the characteristics of the pharmaceutical R&D process. It takes several hundred million dollars to discover, develop, and gain regulatory approval for a new medicine. Absent patent protection, or some equivalent barrier, it would then be the case that imitators could "free ride" on the innovator's FDA approval and duplicate the compound for a small fraction of the originator's costs. Imitation costs in pharmaceuticals are extremely low relative to the innovator's costs for discovering and developing a new compound.

One of the reasons why R&D is so costly in pharmaceuticals is that most new drug candidates fail to reach the market. Failure can result from toxicity, carcinogenicity, manufacturing difficulties, inconvenient dosing characteristics, inadequate efficacy, economic and competitive factors, and various other problems. Furthermore, the full R&D process from synthesis to FDA approval involves undertaking successive trials of increasing size and complexity. Typically, many thousands of compounds are examined in the preclinical period for every one that makes it into human testing. Only 20% of the compounds entering clinical trials survive the development process and gain FDA approval.[13] The preclinical and clinical testing phases generally take more than a decade to complete.[14]

In a recently completed study, DiMasi et al.[15] have examined the average R&D cost for drugs introduced into the market in the late 1990s. Data were collected on R&D costs for a randomly selected sample of 68 investigational drugs from 10 multinational firms. DiMasi et al.[16] found that the representative new product approval incurred out-of-pocket costs of over $400 million.‡ This includes money spent in the discovery, preclinical, and clinical phases as well as an allocation for the cost of failures.

Figure 30.1 shows a breakdown of total R&D costs per approved drug incurred during the preclinical and clinical R&D phases. Expenditures in the clinical period account for roughly 70% of the total out-of-pocket expenditures. This reflects the fact that clinical trials are very expensive on a per patient basis; many drugs must be tested for every one

‡ For an earlier study using the same methodology for 1980s new drug introductions, see Ref. 16.

approved, and drugs that do make it to the final testing phase and FDA submission typically require premarket testing on thousands of patients.

Figure 30.l also shows R&D costs capitalized to the date of marketing at representative cost of capital for the pharmaceutical industry of 11% . The average capitalized R&D cost for a new drug introduction during this period is $302 million, or nearly double the out-of-pocket expenditure. Capital costs are high in this situation because of the long time periods involved in pharmaceutical R&D. More than a decade typically elapses between initial drug synthesis and final FDA approval. Since preclinical expenditures occur several years prior to FDA approval, these costs are subject to greater compounding at the industry cost of capital of 11%.§ Therefore they account for a greater proportion of total capitalized cost compared to total out-of-pocket costs (42 vs. 30%).

R&D costs per new drug approval in the 1990s increased at an annual rate of 7.4% above general inflation when compared to the costs of the 1980s approvals. Major factors driving this increase are the complexity and number of clinical trials that have increased significantly in the 1990s compared to the 1980s. One important factor underlying this trend is the pharmaceutical industry's increased focus on chronic and degenerative diseases; such conditions require larger trial sizes to establish efficacy and longer time periods for observation.

A number of factors could alter the growth pattern for future R&D costs. Emergent technologies may have pro-effects on R&D productivity in the next decade. The mapping of the genome, and related advances in fields such as proteomics and bioinformatics, has led to an abundance of new disease targets. Some industry analysts have hypothesized that these developments may actually cause R&D costs to rise in the short run.[17] The basic reason is that these new technologies require substantial investments upfront, and to date they have generated many disease targets and receptor sites that are not yet well understood. Eventually, expansion in the scientific knowledge base should lead to substantial efficiencies in the R&D process for new pharmaceuticals.

30.3 The Hatch–Waxman Act: Balancing Innovative and Initiative Competition

The patent system is the public policy instrument designed to balance the trade-offs between property rights protection and imitative competition. Without a well-structured

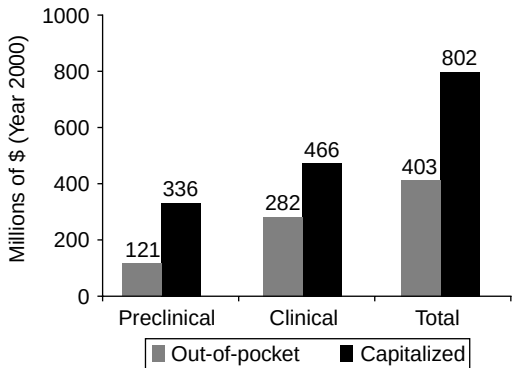


FIGURE 30.1
Out-of-pocket and capitalized costs per approved drug. (From Tufts University Center for the Study of Drug Development.)

§ Capitalization takes account of interest payments and foregone earnings from investments of comparable riskiness during the lengthy R&D investment period for a new drug.

system of global patent protection, neither the research pharmaceutical industry nor the generic industry would be able to grow and prosper, and the rate of new product introductions and patent expirations would decline significantly.

Effective patent life (EPL), defined as patent duration from a product's market launch date, is an important variable influencing R&D incentives in this industry, because it takes many years to recoup the R&D costs and earn a positive return for a typical new drug introduction. Because firms apply for patents at the beginning of the clinical development process, significant patent protection is lost by the length of FDA approval time. This implies a significant reduction in the effective patent life of drugs relative to the nominal life of 20 years.[¶18,19] In light of this, the United States, the European Community and Japan have all enacted patent term restoration laws.

The U.S. law in this regard, the Hatch–Waxman Act, has been in existence since 1984. Hatch–Waxman provides for patent term restoration of time lost during the clinical development and regulatory approval periods, up to a maximum of 5 years additional patent life.[‖] This law also facilitates faster generic product introduction by allowing generic firms to file abbreviated new drug applications, in which generic firms must only demonstrate bioequivalence to the pioneer's products to obtain FDA approval.[**] Prior to the passage of Hatch–Waxman, generic firms had to submit their own proof of a compound's safety and efficacy, as well as show bioequivalence.

Under the act, generic firms can also conduct bioequivalence testing and FDA submissions in the prepatent expiration period so that they can enter the market immediately after patent expiration or invalidation. In addition, the Hatch–Waxman Act allows a generic drug manufacturer to file a "Paragraph 4 Certification," challenging the validity of the patent granted to the branded drug, which in turn triggers up to a 30-month stay while the matter is being litigated. If the challenge is upheld, the generic drug gets a 180-day exclusivity period where no other generic drug is allowed to enter.[20,21]

Grabowski and Vernon[18,19] have investigated the effects of Hatch–Waxman on both generic competition and effective patent lifetimes. Figure 30.2 shows the trends in EPLs by approval year for the new drugs introduced in the first half of the 1990s. This figure indicates that the average EPLs in the 1990s center around an 11- to 12-year range.[*†] The mean for all 126 new drug introductions in the 1990 to 1995 period is 11.7 years with an average Hatch–Waxman extension of 2.33 years. In the last two years of this period, when virtually all of the drugs involve compounds that entered clinical testing after 1984, the average extension is close to 3 years in length. The mode of the frequency distribution of EPLs for this sample of annual new drug introductions is in the interval of 12 to 14 years.

[¶] For data on the trends in effective patent time, see Refs. 18 and 19.

[‖] Title II of the Waxman–Hatch Act provided for partial restoration of the patent time lost during the clinical testing and regulatory approval periods. A formula for patent term restoration was embedded in the law. In particular, new drugs were eligible for an extension in patent life equal to the sum of the NDA regulatory review time plus one half of the IND clinical testing time. The law capped extensions at 5 years and also constrained extensions to a maximum effective patent lifetime of 14 years. Drugs in the pipeline at the time the Act was passed, in September 1984, were limited to a maximum extension of 2 years.

[**] For new drug products with little or no effective patent life, generic firms are prohibited from filing an abbreviated new drug application within the first 5 years of the product life. Most European countries prohibit such filing within the first 10 years of market life.

[*†] This includes any benefits from the international GATT Agreement passed by Congress in 1994, which harmonized U.S. patent laws with foreign countries, including setting the nominal patent life to 20 years from the date of patent application rather than 17 years from the date of patent grant. It does not include any potential benefits of a 6-month extension granted under the FDA Modernization Act in 1997, which can be awarded if the firm does additional testing and gains FDA approval for a pediatric indication.

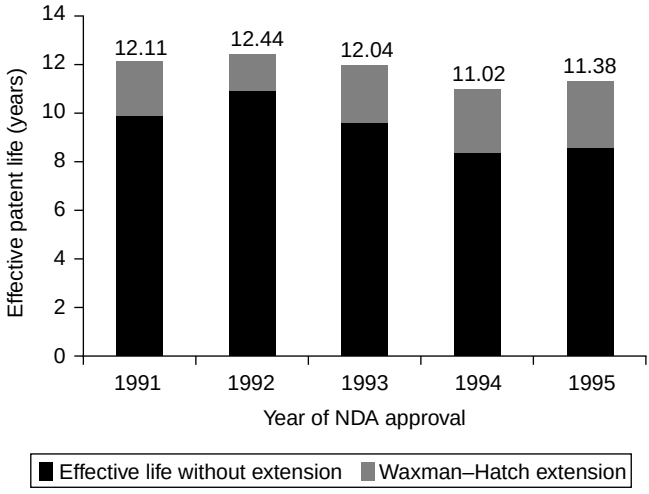


FIGURE 30.2
Effective patent life for 1991–1995 NCEs.

30.4 Generic Competition Since the Act

In contrast with new product introductions, the development costs of generic compounds are relatively modest. In the United States, since the passage of the 1984 Hatch–Waxman Act, generic products need only demonstrate that they are bioequivalent to the pioneering brand to receive FDA approval. Generic firms can file an Abbreviated New Drug Application (ANDA), a process that takes only a few years and typically costs a few million dollars.[22,23] Also, the probability of success is very high, as reflected by the fact that many generic firms file to receive FDA approval and enter the marker within a short window of opportunity around patent expiration of the pioneer brand.

A distinctive pattern of competitive behavior for generic and brand name firms has emerged in the wake of the 1984 act. First, commercially significant products experienced a large number of generic entrants within a short time after patent expiration, in sharp contrast to what occurred in the pre-1984 period. Also, in the post-1984 period, a strong positive relationship between the size of the market and the number of generic competitors can be noted in accordance with expectations from economic theory.

Second, generic drugs exhibited a high degree of price competition after 1984. The initial generic product entered the market at a significant discount compared to the brand name product, and this discount grew larger as the number of generic competitors for a particular brand name product increased over time. In a study of commercially significant products from 1984 to 1989, Grabowski and Vernon[24] found that generic prices averaged 61% of the brand name product during the first month of generic competition. This declined to 37% by 2 years after entry.

Third, a more rapid rate of sales erosion in brand name products was observed in the case of more recent patent expirations. This is especially so for the top selling drug products which attract the most intensive generic competition. Three recent cases illustrate this phenomenon. Generic enalapril, launched in August 2000 as a substitute for the brand name drug Vasotec®, obtained 66.4% of new prescriptions just 4 weeks after its launch.

Generic lisinopril, launched in May 2001 against the brand name drugs Zestril® and Prinivil, acquired 84% of new prescriptions in 4 weeks. Generic fluoxetine, launched in August 2001 against the brand name drug Prozac®, acquired 74.9% of new prescriptions in 4 weeks.*‡

Congressional Budget Office[22] has also done an analysis of the economic effects of Hatch–Waxman. As in Grabowski, and Vernon's analysis, they found that generic competition has been a powerful force for price competition since 1984. The CBO estimated annual savings of $8 to $10 billion to consumers by the mid-1990s. In terms of R&D incentives, however, they found that Hatch–Waxman had negative effects on the expected returns on R&D. In this regard, they estimated that the act, together with the increased demand side incentives promulgated by managed care organizations to utilize generic products in the 1990s, has resulted in steadily accelerating erosion of pioneer-brand's sales in the period after generic entry.

Overall, price competition and generic utilization have increased dramatically since the Hatch–Waxman Act was passed. In 1984, generic products accounted for approximately 14% of all prescriptions. By 2002, the figure was 51%.*§ The growth of managed care and other related demand-side changes have been important factors underlying this rapid increase in generic usage. However, the passage of the 1984 act played a critical role by relaxing the regulatory hurdles for generic firms and facilitating higher levels of generic entry.

30.5 Are the Innovations and Imitation Costs of New Biotech Entities Different?

Most of the analyses of R&D costs for new dug entities and their generic imitators have focused on small molecule NCES. This reflects the relative youth of the biotech industry. New biologic entities were first introduced in the 1980s. By 1994, only 29 new biological entities had been introduced into the U.S. market, but this number has increased dramatically since then. In this regard, 41 new biological introductions occurred between 1995 and 2001.

The newest R&D cost study by DiMasi et al.[15] includes seven biotech compounds, in a sample of 69 entities, for which data were obtained from major pharmaceutical and biopharmaceutical firms. While this sample of biological entities is too small to be representative of all biotech development, the clinical phase costs in the DiMasi et al. study were similar for the biotech and pharmaceutical projects.

As discussed earlier, capitalized R&D costs per new drug introductions are influenced by a number of factors. These include out-of-pocket costs at the preclinical and clinical phase, the probability of success for new drug candidates at different stages of the R&D process, and the length of time that it takes to move through all the stages of the R&D process and gain FDA approval. Recent studies of the probability of success and the length of the R&D process for biotech drugs indicate a convergence in these parameters toward the values observed for small molecule pharmaceuticals.

*‡ This issue also has been examined using a broad sample of products over the period 1984 to 1997.[18,19]
*§ PhRMA, Pharmaceutical Industry Profile 2003, Available at www.pharma.org

Two initial studies found that success rates for biotech drugs were substantially higher than success rates for NCEs.[25,26] In particular, both studies projected success rates for biopharmaceuticals in excess of 50%. However, a basic assumption implicit in the methodology of both studies is that success rates for biotech drugs that entered development in the late 1980s and early 1990s are the same as for the biotech drugs that entered development in the early to mid-1980s. This is a very strong, and potentially hazardous, assumption given that 90% of the drugs in their samples were still under active testing.

Subsequently, Gosse et al.[27] analyzed a comprehensive sample of U.S. biopharmaceutical drugs and compared the success rates of older and newer biotech entities. They found dramatic differences in the time pattern of success rates observed for early vs. later biotech drug cohorts. In particular, for the investigational new drugs (INDs) filed in the early 1980s, the success rate for new recombinant entities is 38%. For the INDs filed during the late 1980s the success rate was only 10% based on approvals to date (i.e., 6 years after testing). At a comparable point in time, the new recombinant entities of the early 1980s had a success rate of 26%. In fact, the success curve of the recent recombinant entities more closely resembles that of NCEs rather than that for the early biological entities.

This result is consistent with the history of biotech research in the United States. The first biological entities introduced into the market were naturally occurring proteins that replaced purified nonrecombinant formulations already in general use as established therapies (e.g., insulin and human growth hormone). It is reasonable to expect that recombinant versions of established therapies would have high success rates, once the technology to manufacture these products was proven. Other earlier targets for biotechnology were naturally occurring proteins with well-known and defined physiologic activity (e.g., erythropoietin and filgrastim). As the biotech drugs moved to targets for which limited knowledge existed about clinical and pharmacological profiles, it is reasonable to expect that success rates would fall back toward those of conventional drug entities.

The prospect of a long and uncertain discovery and development period for a new drug is another factor affecting costs and risks in the drug R&D process. The longer the development and approval process, the higher the interest costs, opportunity costs, and the overall capitalized R&D costs of a new drug introduction. Recently, Janice Reichert of the Tufts University Center for the Study of Drug Development has done a historical analysis of clinical development time for successive cohorts of new biopharmaceuticals. She found that recently introduced biopharmaceuticals had much longer clinical development times than earlier introductions. In particular, the cohort of 2000–2001 new biopharmaceutical introductions had a total clinical development time (including FDA approval) of 86 months, vs. 53.2 months for 1982 to 1989 biopharmaceutical introductions.*¶

Hence, the experience with respect to development times parallels the experience observed with respect to success rates. In particular, there has been a convergence in clinical trial period times observed for new biological and NCEs. Of course, the biotech industry is still in the early stages of evolution. It may eventually produce higher success rates and shorter development times as a result of new technologies currently emerging in the discovery period. However, the best evidence at this time shows that biopharmaceuticals, like NCEs, are subject to very high rates of attrition and long gestation periods in the clinical development stage.

One aspect in which biopharmaceuticals may differ from small molecule NCEs concerns the ease of generic entry when patents expire. To date, there have only been a few patent

*¶ These data were obtained in April 2002 by Mom Janice Reichert at the Tufts University Center for Study of Drug Development in Boston, Massachusetts. For further trends and analyses in this regard, see the Tufts Center's *Inc pact Report*, Vol. 3, Nov/Dec 2001.

expirations involving biopharmaceuticals. One case in which there has been entry after patent expiration is that of human growth hormone. However, all entry to date has been by other big pharmaceutical firms that have had experience in supplying this product in Europe and Japan (Pharmacia, Novo Nordisk and Ares Serono). There are greater hurdles in manufacturing biopharmaceuticals at an efficient scale compared to NCEs, and in addition there are greater regulatory requirements for biologicals associated with the manufacturing process.[28] These factors may moderate the degree of imitative competition for biopharmaceuticals compared to small molecule chemical entities. Whether or not this is the case will become more apparent when some of the commercially important biopharmaceuticals are subject to patent expiration and potential generic competition in the coming decade.

30.6 Returns on R&D for New Drug Introductions

Grabowski et al.[29] have examined the distribution of returns for new drug introductions.*[||12] This work builds directly on the R&D cost analysis of DiMasi et al., and considers the sales and net revenues realized over the product life of new drug introductions during the 1970s, 1980s, and 1990s. One finding of this work is that the distribution of returns to new drug introductions is highly variable, noting another source of risks for firms developing new drug introductions.

Figure 30.3 shows the distribution of present values of net revenues (revenues net of production and distribution costs but gross of R&D investments outlays) for new drug introductions between 1990 and 1994. The distribution shows very strong skewness. Roughly one half of the overall present value from this sample of 118 compounds is accounted for by the top-ranked decile of new drug introductions. The top decile has an estimated after-tax present value that is more than five times the present value of average after-tax R&D costs per approved introduction. Furthermore, only the top three deciles of new drug introductions have present values that exceed average R&D costs.

A major factor underlying the skewed distribution observed in Figure 30.3 is the level of sales realized by new drug introductions. A few drugs achieve peak sales of several billion dollars and account for a large share of overall revenues. At the other end of the distribution, many compounds achieve peak sales only in the tens of millions of dollars and fail to provide a positive return on investment.

Grabowski and Vernon have investigated other periods and time cohorts of new introductions and found that they are characterized by similar patterns.[29]

These returns to R&D analyses confirm that the search for blockbuster drugs is what drives the R&D process in pharmaceuticals. The median new drug introduction does not cover average R&D costs (including allocations for the cost of discovery and the candidates that fall by the wayside). A few top-selling drugs are key in terms of achieving economic success in pharmaceutical R&D over the long run. The large fixed costs of pharmaceutical development and the skewed distribution of outcomes help to explain the clustering of biotech firms at the research stage of the R&D process and the large number of alliances between biotech and big pharmaceutical firms at the development and marketing stages.

*[||] For earlier studies of new drug introductions in the 1970s and 1980s, see Returns to R&D on new drug introductions in the 1980s, *J. Health Econ*., 13, 383–406, 1994; and A new look at the returns and risks to pharmaceutical R&D, *Manage. Sci*., 36, 804–821, 1990.

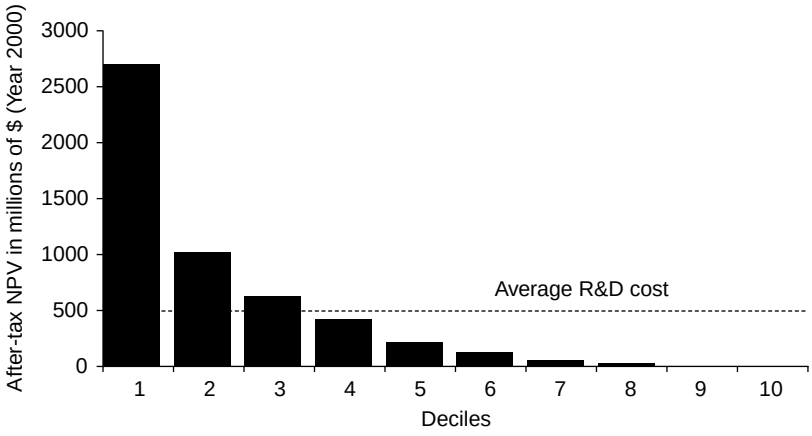


FIGURE 30.3
Present values of net revenues by decile: 1990–1994 NCEs.

In Figure 30.4, the distribution of worldwide sales in 2000 is presented for 30 new biological entities introduced into the U.S. market between 1982 and 1994. All the compounds had been in the market at least 7 years, and had progressed beyond the initial rapid growth phase of their life cycle. The sales data presented in Figure 30.4 indicate that new biopharmaceuticals also exhibit a high degree of skewness, similar to the much larger cohort of new drug introductions.

The high degree of skewness in the outcomes of pharmaceutical R&D projects indicates that there are substantial risks in this endeavor, both for big pharmaceutical firms as well as smaller biotech enterprises. This reflects the dynamic nature of the R&D process, the long time periods from the start of a project to market approval, the unpredictability of clinical trials, as well as regulatory and competitive uncertainties.

Even though many big pharmaceutical firms spend billions of dollars per year on a diversified portfolio of in-house and out-sourced projects, this does not guarantee a stable set of outcomes. If a firm invests in a large diversified portfolio of projects that are more tightly clustered around the mean return (e.g., the so-called normal or Gaussian distribution), we expect that returns can be predicted with sonic confidence. When returns are highly skewed, however, individual companies experience highly volatile outcomes even when they invest in large numbers of independent projects.

To illustrate this point, Grabowski and Vernon examined the new product sales for the U.S. drug companies that spent between $300 and $500 million on their global R&D in the mid-1980s (the top tier group in that period). They found that subsequent new product sales emanating from these R&D efforts varied between $100 million and $3 billion (after 7 years of marker life).[18,19]

Finally, it is important to note that the distribution of outcomes from pharmaceutical R&D projects has similar characteristics to many other innovation samples including venture capital funding of high tech start-ups. In this regard, Scherer et al.[30] have examined the size distribution of profits from investments in innovation projects using a diverse set of data samples. Their analysis included two large samples of high-technology venture capital investments as well as a comprehensive sample of venture backed start-up firms that had their initial public offering in the mid-1980s. A common finding was that the size distribution of profit returns from technological innovation is strongly skewed to the right. As in the case of new drug introductions, the most profitable cases contribute a disproportionate fraction of the total profits from innovation.

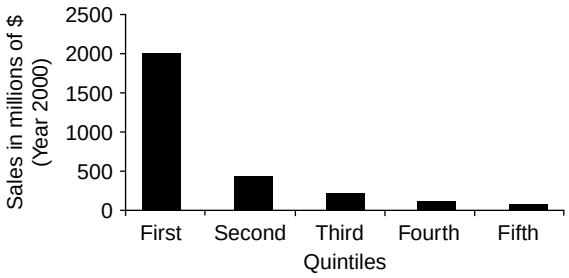


FIGURE 30.4
New biotech introductions 1982–1994 worldwide sales in 2000.

TABLE 30.1
Returns Distribution for Selective Innovation Samples

Date Set	From Top Decile (%)
Venture capita (start-ups)	62
Horsley–Keough (start-ups)	59
1980s IPOs — (1995 value)	62
1990s New drugs (Grabowski–Vernon)	52

Table 30.1 summarizes the results from four data sets employed in Scherer's analysis. The first two datasets, assembled by Venture Capital Incorporated and Horsley Keough Associates, involve an analysis of several hundred venture capital firm investments in high tech start-up companies. Scherer's analysis indicates that roughly 60% of the returns, measured at the time of the final distributions to investors, are realized by the top decile of venture capital projects. At the same time, roughly half of the projects in these samples failed to earn positive returns. Similarly, an analysis of the stock market performance of the universe of high tech companies that were public in the mid-1980s found that the top decile of companies realized 62% of the samples total market value 10 years later. The corresponding value for Grabowski and Vernon's sample of 1990 to 1994 new drug introductions is 52%. Hence the three samples of risky, high-tech start-up companies exhibit skewed distributions of returns comparable to the pharmaceutical industry.

30.7 Conclusions and Policy Considerations

Economic analyses of the R&D process in pharmaceuticals indicate that it is a very costly and risky process, even for large established firms. Most compounds in the R&D pipeline never reach the marketplace. The development process is time-consuming and expensive, and the distribution of profits among those products that are marketed is highly skewed. A few blockbuster successes cover the losses on many other R&D investment projects. An important implication for public policy is that reimbursement, regulatory or patent policies that target the returns to the largest selling pharmaceuticals can have significant adverse consequences for R&D incentives in this industry.[31]

Many compounds in the top decile of the returns distribution involve the first mover, or other early entrants, in a new therapeutic class. The family of medicines in a given

therapeutic class passes through a well-delineated life cycle. There is dynamic competition involving breakthrough, as well as incremental advances, among the branded products within that class. This dynamic competition, in turn, produces substantial consumer surplus and social returns. When the patents for established products expire, consumers also benefit from imitative competition from generic entrants, which provide social benefits in terms of significantly lower prices.

Patents play a critical role for both the level of R&D investment in pharmaceuticals and the timing of generic competition. The Hatch–Waxman Act was designed to balance the trade-off between these two objectives. In particular, it sought to produce patent lifetimes sufficient to encourage increased levels of R&D investment by innovators while promoting intense price competition by easing entry regulations to generics when patents expire. The degree of generic competition has increased dramatically since the 1984 act was passed, with more than half of all new prescriptions accounted for by generic products in 2002.

The Watch–Waxman Act has fostered a vigorous generic industry with substantial benefits to consumers from price reductions. However, the CBO[21] analysis of the act found that from the perspective of R&D returns, the much more rapid loss of sales in the period after patent expiration has dominated the patent term restoration aspects of the law. In particular, using Grabowski and Vernon's analysis of R&D returns as their conceptual framework, they estimated 12% lower expected after-tax profits from R&D for the mean new drug compound as a consequence of the 1984 act. While the mean compound is still moderately profitable in their analysis, the increased generic competition since 1984 can have adverse R&D incentives for compounds of above average riskiness or ones with shorter than average effective patent life.

Overall, Hatch–Waxman has provided a relatively balanced approach to the trade-offs between pharmaceutical R&D and generic competition. Improvements on the margin could be considered by policy makers, such as a longer minimum exclusivity period before an ANDA could be filed for new drag introductions (currently 5 years in the United States but longer in Europe and Japan).†* Nevertheless, the law has provided a reasonably well structured system of incentives for both innovative and generic firms. Both R&D investments and generic utilization have increased dramatically in the period since the passage, consistent with the objectives of the act. Some groups have suggested that Congress consider significantly collecting or eliminating the patent restoration aspects of the law to further increase generic competition in pharmaceuticals.†† Given the critical role that patents and effective patent life play in terms of R&D incentives for this industry, this would not appear to be a desirable course of action on social welfare grounds.

References

1. *The Georgetown Public Policy Review*, 8, 7–24, 2003; *Science & Cents*: *Exploring the Economics of Biotechnology*, in *Proceedings of a Conference, 2002*, Sponsored by the Federal Reserve Bank of Dallas, Ducca, J.V. and Yucel, M.K., Eds., Federal Reserve Bank of Dallas, 2003, pp. 87–104.
2. Grilliches, L., The search for R&D spillovers, *Scand. J. Econ.*, 94(Suppl.), 29–47, 1992.
3. Scherer, F.M., Industrial market structure and economic performance, Rand McNally, Chicago, 1980.

†* Grabowski and Vernon[18,19] found that relatively few NCEs are marketed with effective parent lifetimes of less than 10 years.

†† See, for example, National Institute for Health Care Management Foundation, August, Prescription drugs and intellectual property protection, NICHM Foundation, *Issue Brief* (Washington, DC).

4. Cutler, D.M. and McClellan, M., Is technological change in medicine worth it? *Health Affairs*, (Sept/Oct.), 20, 11–29, 2001.
5. Triplett, J.E., Ed., *Measuring the Price of Medical Treatments*, Brookings Institution, Washington, DC, 1999.
6. Towse, A., Ed., *Industrial Policy and the Pharmaceutical Industry*, Office of Health Economics, London, 1995.
7. Levin, R.D., Klevorick, A., Nelson, R.R., and Winter, S.G., Appropriating the returns from industrial research and development, *Brookings Papers Econ. Activ*., 783–820, 1987.
8. Cohen, S.M., Nelson, R.R., and Walsh, J.P., Appropriability conditions and why firms patent and why they do not in the American manufacturing sector, Working Paper, Carnegie–Mellon University, Pittsburgh, 1997.
9. Taylor, C.T. and Silberston, Z.A., *The Economic Impact of the Patent System*, Cambridge University Press, Cambridge, U.K, 1973.
10. Silberston, Z.A., *The Economic Importance of Patents*, The Common Law Institute of Intellectual Property, London, 1987.
11. Mansfield, E., Patents and innovation: an empirical study, *Manage. Sci*., 32, 175, 1986.
12. Baker, D., A haywire parent system picks your pocket at the drug store, *Los Angeles Times*, March 4, 2003.
13. DiMasi, J.A., Success rates for new drugs entering clinical testing in the United States, *Clin. Pharmacol. Therap*., 58, 1–14, 1995; DiMasi, J.A., Trends in drug development costs, times and risks, *Drug Inform. J*., 29, 375–384, 1995.
14. Kaitin, K.I. and DiMasi, J.A., Measuring the pace of new drug development in the user fee era, *Drug Inf. J*., 34, 673–680, 2000.
15. DiMasi, J.A., Hansen, R.W., and Grabowski, H.G., The price of innovation: new estimates of drug development costs, *J. Health Econ*., 22, 151–185, 2003.
16. DiMasi, J.A., Hansen, R.W., and Grabowski, H.G., The cost of innovation in the pharmaceutical industry, *J. Health Econ*., 10, 107–129, 1991.
17. Lehman Brothers, The fruits of genomics: drug pipelines face indigestion until the new biology ripens, New York, January, 2001.
18. Grabowski, H.G. and Vernon, J., The distribution of sales from pharmaceutical innovation, *Pharmco. Econ*., 18(Suppl. I), 21–32, 2000.
19. Grabowski, H.G. and Vernon, J., Effective patent life in pharmaceuticals, *Int. J. Technol. Manage*., 19, 98–120, 2000.
20. Department of Health and Human Services, Food and Drug Administration (DHHS), and Center For Drug Evaluation and Research, Guidance for industry: court decisions, ANDA approvals, and 180-day generic exclusivity under the Hatch–Waxman amendment to the Federal Food, Drug, and Cosmetic Act, DHHS, Washington, DC, March, 2000.
21. Federal Trade Commission (FTC), Generic drug entry prior to patent expiration, An FTC Study, FTC, Washington, DC, July, 2002.
22. Congressional Budget Office (CBO), How increased competition from generic drugs has affected policy and returns in the pharmaceutical industry, U.S. Government Printing Office, Washington, DC, 1998.
23. Reiffen, D. and Ward, M.R., Generic drug industry dynamics, Federal Trade Commission Working Paper 248, February, 2002.
24. Grabowski, H.G. and Vernon, J., Brand loyalty entry and price competition after the 1984 act, *J. Law Economet*., 35, 331–350, 1992.
25. Bienz-Tadmor, B., Cerbo, P.A.D., Tadmor, G., and Lasagna, L., Biopharmaceuticals and conventional drugs clinical success rates, *Biotechnology*, 10, 521–525, 1992.
26. Struck, M.M., Biopharmaceutical R&D success rates and development times, *Biotechnology*, 12, 674–677, 1994.
27. Gosse, M.E., Manocchia, M., and Nelson, T.F., Overview of US Pharmaceutical Development, 1980–1994, Tufts University Center for the Study of Drug Development, May, 1996.
28. Grabowski, H.G. and Vernon, J., The search for new vaccines: the errors of the vaccines for children program, American Enterprise Institute, Washington, 1994.
29. Grabowski, H.G., Vernon, J., and DiMasi, J., Returns on R&D for new drug introductions in the 1990s, *Pharmco. Econ*., 20(Suppl 3.), 11–29, 2002.

30. Scherer, F.M., Harhoff, D., and Kukies, J., Uncertainty and the size distribution of rewards from innovation, *J. Evolut. Econ*., 10, 175–200, 2000.
31. Grabowski, H.G. and Vernon, J., Prospects for returns to pharmaceutical R&D under health care reform, in *Competitive Strategies in the Pharmaceutical Industry*, Robert Helms, Ed., AEI Press, Washington, 1996.

31

The Pharmaceutical Revolution: Drug Discovery and Development

John C. Somberg

CONTENTS

31.1 Introduction

The pharmaceutical industry is undergoing revolutionary change with the impact of new technologies, biotechnology, genomics, proteomics, and the effects of the most profound change in Medicare in the last 40 years — the prescription drug benefit. The forces of technology and government are like fast-running rivers, forever changing the topography of the landscape. Change is a part of the human condition that we often fail to notice. During the last 100 years, the drug development process has undergone considerable change, but all that has gone before will pale in comparison with the dramatic, new information era, which will markedly alter the environment we work in and the drug development process.

Even the most powerful and financially stable companies engaged in drug discovery and development need to recognize the forces of change. The evolution of the computing age gave IBM (Armonk, NY, U.S.A.) the opportunity to expand and alter its business from

analog systems and adding machines to punch cards and then to complex computer systems. As the computer age developed, IBM led the way with the innovative personal computer. The lead was then lost by IBM when "software" became the core of the information age, along with the chips that permit the exponential growth in machine computing performance. Thus, Microsoft, just a concept 30 years ago, is more dominant in today's information age than IBM. IBM remains a leader in technology advances, in new, fundamental patents, and a formidable presence in the information technology (IT) departments of large corporations. It kept up in technology, but its management failed to perceive the salient change in the information age from large computers to small, and then, the importance of the "software" that controls information processing, analysis, and communication.

Drug development is in an analogous situation. We have seen an evolution from the age of botanicals to the age of chemicals and synthetic discovery and, now, to the age of biotechnology and gene manipulation. Each area can still grow, but the shift in direction is fundamental to scientific development. Taking these major changes into account, drug discovery and development will be fundamentally altered by the information age. The concept of the information age had been coined by Alvin Toffler[1] and was correctly perceived to be revolutionary in its effects on society. The agricultural and the industrial revolutions brought about fundamental changes in society, as will the information age. In his book entitled *Future Shock*, Toffler describes an era when the pace of change in modern life is so great as to disenfranchise individuals from the process that society is undergoing. Although this is a real problem for society — a problem with political dimensions — failure of our institutions and of our corporate structures to adjust will bring considerable societal and economic disruption. For these reasons, an understanding of the evolution of drug discovery and development and how this evolution will be affected by the information age is essential for those working in this area.

31.2 Drugs from Plants

From the earliest times, dating at least to the hunter and gatherer societies, humans made use of herbals. Whether as foods, items of religious significance, or medication cannot be clearly discerned. As civilization progressed, remedies from plants developed further. Earliest folklore relates stories of plant medicinals. The Bible contains passages alluding to medicinal herbs and plants. In fact, all the major religions discuss plant remedies as part of their sacred heritage.

There are many stories in pharmacology relating to the use of medicinal plants and the work of herbalists in the early discovery of drugs. One example is the story of William Withering, the physician from Birmingham, England, who on his charity rounds in Shropshire saw that an herbal potion was used to treat a woman with dropsy (congestive heart failure, CHF) who then showed improvement. Withering's botanical training in Edinburgh permitted him to identify the probable active ingredient, the leaf of the foxglove plant. After 10 years of clinical experimentation, he developed a series of case studies explaining the dose range: from minimal effective dose to toxicity. He categorized the adverse side-effect profile of the digitalis leaf and its potentially life-threatening toxicities. He noted the adverse outcomes and carefully chronicled the conditions in which the drug was most useful. Although he thought the agent increased urine volume and, therefore, had diuretic properties, he commented in his thesis that the drug had a powerful action on the motion of the heart and, thus, recognized its cardiotonic action years before this was actually proven. Withering was a masterful botanist (he chronicled the plants of Great

Britain later in his life). He was an exemplary clinical pharmacologist and demonstrated the best in botanical drug discovery and testing given the skills of his day. But Withering's observations may not be unique. The effect of the foxglove plant on disease was known to be part of European plant folklore. The use of these glycoside-yielding plants and the use of the skin of the toad for medicinal purposes go back to ancient Egypt and are mentioned in Chinese herbal writings. Confucius talks of glycoside plants for edematous states, and cardiac glycosides are a significant component of Chinese herbal medications. Although Withering's observations were a defining moment for modern medicine, botanicals of similar action were used for more than 2000 years.

Clearly, botanicals have been an important component to therapeutic advances. Whether we are discussing digitalis, atropine, or any number of other drugs, plants have contributed much to drug discovery. Using quinidine in atrial fibrillation or quinine to treat malaria are other examples of the importance botanicals have played in therapeutics. In fact, in the 1700s and 1800s, botanicals were the only source of drugs for development. The antiinfective agents have depended on extracts from molds and fungus for a very long time. Recent therapies are derived from nature with some chemical modifications to improve activity.

We often think of the age of botanicals as one that has gone by. Indeed, it was the first step in the field of drug discovery and development, but one that continues to this day to play a major role. Although reserpine was used for a 1000 years in India and parts of China, it was only in the 1950s that it was purified and used as an effective antihypertensive agent. The recent use of taxol in oncology is an example of a botanical that was in very short supply. Until a synthetic pathway for commercial production was developed, the bark of the yew tree became a very valuable commodity and caused the yew tree to be endangered. The cardiac antirejection drug sirolimus (Rapamune, Wyeth Laboratories, Collegeville, PA, U.S.A.) is an even more recent example of a soil sample taken years earlier from an atoll in the Pacific Ocean that has led to a major new therapy. The therapy is used both as an antirejection drug in transplantation surgery as well as a coating for coronary artery stents to prevent restenosis.

There was a trend for large companies to form alliances with botanical gardens and countries and explore to find new therapies. In fact, a company called Shaman Pharmaceuticals (South San Francisco, CA. U.S.A.) was founded with a corporate purpose to discover and develop pharmaceuticals from botanical sources based on the folklore of "medicine men." However, these ventures and alliances were not quick to deliver new therapies, and the impetus to continue this phase of development is fading. This is unfortunate because this type of discovery has proven so fruitful in the past, and nothing now known should discourage this type of exploration. In fact, the one discouraging factor is the rapid "drop off" in the planet's biodiversity that has been caused by the relentless expansion of human development with the emphasis on corporate agriculture, the clearing of the Amazon jungle, international travel, contaminating the biomass, and the worldwide, rampant growth of "nonnative species."

To successfully deal with the challenge of drug development from botanicals requires the application of the most modern techniques, the most important of which may well be those related to handling the vast amount of information that can be collected. Clearly, computer applications for exploration of the plant world include automating processes for analysis, and chemical categorizing with innovative storage, organization, and retrieval, which will all be required to make the drug discovery process effective. The systematic computerization of knowledge in ethnobotany and pharmacognosy, with emphasis on plant categorization across primitive societies, will be helpful in sustaining the discovery process. Using sophisticated computer techniques to look for similarities in medicinal plant use among primitive peoples to ascertain potentially useful observations can greatly

aid the ethnobotanist. Hopefully, these computerized techniques will replace the hundreds, if not thousands, of years that are needed for serendipitous observations such as that made by William Withering 200 years ago, which led to the introduction of the digitalis glycosides in clinical medicine.

31.3 Synthetic Drug Discovery

For the last 75 years, most of the new molecules have come from synthetic chemistry. In cardiology, the β-blockers and calcium channel blockers have revolutionized cardiovascular therapeutics. β-Agonists in respiratory therapy and H_2 antagonists in gastrointestinal ulcer disease therapy are a few examples of the work of the synthetic chemist that has greatly changed our treatment of patients. The proven model for finding a useful transmitter in a physiological system is to find a receptor to which the transmitter reacts and, then, modifying the agonist structure to find a specific antagonist.

As new receptors and new physiological systems are revealed, the synthetic chemist will surely be making considerable contributions to the field of drug discovery. The process is ongoing. For instance, as the role of the endothelium becomes better understood, its impact on pharmaceutical research greatly expands. What was once called endothelial-derived relaxing factor (EDRF) has been characterized as a locally released gas, nitric oxide. Studies on endothelium function have found endogenous substances involved in the modulation of vasodilation and vasoconstriction at the local endothelial level. There are endogenous substances opposing the vasodilating properties of nitric oxide. Endothelin is one of these transmitters, and the development of specific endothelin antagonists is an exciting new field. Whether these endothelin antagonists will be effective therapies in angina, hypertension, or CHF remains to be determined, but the process shows that the synthetic discovery of drugs, combined with physiological transmitter research, is still spawning drugs of great potential.

Even here, with well-established approaches, we see the influence of the information age. Employing computers to determine receptor structure and, thus, possible receptive blockers has become a useful tool in the drug discovery process. Computer-assisted drug synthesis has great potential. The revolution in this aspect of synthetic chemistry is analogous to the revolution that computers caused in the animation industry. Where once dozens of artists were necessary, computers have now replaced them, creating "life-like" animations that were not previously feasible. The same type of revolution will occur in the chemical drug synthetic industry.

Besides design, computerization can readily be applied to categorizations and synthetic pathways. I believe the application of computer sciences to chemistry will lead to considerable advances in this field. The application of computers to the steps beyond modeling systems — to identifying chemical structures and automatically developing synthetic approaches — will have a considerable effect. Synthetic antagonists with optimum potency can be developed from a host of chemical possibilities. With a heightened receptor selectivity and potency, the increased yield of these procedures will be noticeable. What industry must target are improved therapies as well as new and novel therapies. A more potent, less-toxic agent can be as useful and financially rewarding as a blockbuster.

However, the information age applications to synthetic modeling will be inherently limited unless we can improve our screening techniques. For many years, I have given considerable thought to the link between drug synthesis, discovery, and development.

Almost 20 years ago, I had the good fortune to visit Jansen Pharmaceuticals (Breeze, Belgium) and discuss the drug discovery process with the late Paul Jansen, a genius in the field of drug development. I was most impressed with his grasp of chemistry, his diverse interests, and his unparalleled success in the discovery of novel entities. Jansen was a chemist looking for novel compounds that could then be assessed to find biological activity. A new, promising compound would be processed through hundreds of models, looking for possible pharmacologic activity. The question arose about the ability to screen for biological activity. This is a critical linkage point in the discovery and development process, one to which the great potential of the information age can be effectively applied.

Jansen's approach fascinated me, but to this day that visit caused some abiding concerns. Because I have been interested in the field of antiarrhythmic drug development for many years and have participated in all stages of drug development from chemical synthesis to the clinical arena of programed electrical-stimulation studies, I was particularly interested in Jansen's approach as it applied to antiarrhythmic pharmacology. I had been working on lorcainide, a drug Jansen developed at Bersa-Chem (Jansen, Belgium), and wondered how this compound was discovered and how it compared to other agents screened. Jansen employed a costly dog model of premature ventricular contraction (PVC) suppression postcoronary artery ligation. Lorcainide — a Vaughan–Williams Ic antiarrhythmic agent — was a sodium channel blocker. With this profile, lorcainide was predicted to be effective in the PVC-suppression model. But PVC suppression and the Ic agents have not shown prolongation of life in the postmyocardial infarction (MI) studies.[2] The type III agents appear to be most effective clinically, although not shown to prolong life in randomized controlled trials (including the European Myocardial Infarction Amiodarone Trial and the Canadian Amiodarone Myocardial Infarction Trial). However, a metaanalysis of drugs of the type III variety and, specifically, amiodarone has suggested them to be far more effective with much less proarrhythmia than the sodium channel blockers. This leads one to consider what the effect would be of the clinically valuable agent amiodarone in the screening model that Jansen was using to pick out his antiarrhythmic to go into clinical development. In fact, the records at Bersa were so accurate that the scientists in that department could look up the results in a few minutes and describe the actions of other known antiarrhythmic agents in the drug model. The answer they gave was that amiodarone was much less effective and, in fact, hardly effective at all in the model in which lorcainide was extremely effective. It is no wonder that the pharmaceutical industry in the 1980s found a host of Ic agents (flecainide, encainide, lorcainide, propafenone, indecainide, and ethmozine, for instance), because that is what their assays were best at picking up as active agents.

The model, therefore, is critically important in the drug discovery process, and it will often determine development. We could synthesize thousands of compounds, then select a few for development that may not be optimum for therapy. These agents, however, would fit the characteristics being sought by the model employed in the screening process. This is a major problem that has not been given enough consideration. We can only think of the possibility that there may be hundreds, if not thousands, of compounds buried in analytical hoppers such as Jansen's Bersa research establishment that could have been extremely useful but were discarded because they were not identified as biologically active in an inherently flawed screening model.

In addition to the models used in drug screening, there is a fundamental difference in discovery between mass screening and receptor-targeted research. The latter has proven more successful in the past decade, but some major advances have come out of pure chemistry and follow-up screening to determine biological activity. Can the revolution of the information age and computer sciences be applied to synthesis and screening? These questions will challenge us in the coming years. I believe a revolution will occur in this area.

Synthesis on a grand scale will be tied to automated, focused, biological-activity screening that will permit the evaluation of tens of thousands of molecules on a daily basis. Clearly, how we screen will determine the validity of this approach.

Although we are in transition from the age of synthetic chemistry to the age of biotechnology and gene manipulation, synthetic discovery will still play a major role in advancing the therapeutic armamentarium.

31.4 Biotechnology, Genomics, and Proteomics

The area of biotechnology and gene manipulation is in its early stages but has already made a considerable impact. The largest companies are busy positioning themselves by acquiring or "joint venturing," with biotechnology companies, which are often small, start-up enterprises. The pharmaceutical companies are undertaking these acquisitions so that they will be prepared to benefit from the coming revolution in biotechnology and gene manipulation. Biotechnology has not advanced as rapidly as predicted. The science has made tremendous strides, but a number of factors have limited the advances and commercialization.

The scaleup and commercialization of biotechnology processes is limited by expense and the difficulties that are imposed technologically. The first generation of compounds can be effective, as demonstrated by growth hormone and recombinant tissue plasminogen activator (rTPA). However, there have been major failures, such as the antibodies to counteract the effects of septic shock. Whereas science permits the creation of drugs for evaluating, the compounds themselves may not be effective. This dichotomy stems from our imperfect knowledge of the pathophysiology of disease states, such as is the case with Gram-negative sepsis and shock. Another problem revolves around a constantly changing target, such as the AIDS virus.

Genentech (South San Francisco, CA, U.S.A.) and Amgen (Thousand Oaks, CA, U.S.A.) have been successful in bringing drugs to the marketplace, but even these companies have struggled to remain viable and to maintain adequate cash flow for undertaking the research and development for the next generation of products. The hundreds, if not thousands, of smaller companies may not fare as well, and it is safe to predict that only a small fraction will indeed find a successful product. Besides the discrepancy between the ability to make a compound and its clinical efficacy lie the problems of corporate capitalization and effective drug development. The mergers of biotech concerns and the established pharmaceutical industry will go beyond improved capitalization and will bring more expertise in drug development and the regulatory approval process to this fledgling industry. But there are further impediments to success. Many of the products of biotechnology synthesis are proteins that are not orally active. A major area of research is going to be to convert the intravenously active compounds into ones with a facilitated means of delivery. Novel drug delivery systems to overcome the problems of lack of oral activity will be crucial. Carrier molecules, topical transport enhancers, and nasal absorption enhancers and methodologies are but a few of the possible solutions to the drug delivery problem that considerably hampers the biotechnology field. Another approach has been the development of chemical molecules that have similar, key, structural elements that may permit the chemical compound to act like the protein molecule. If this is possible, we may find ourselves using the tools of biotechnology to enhance the drug discovery process through chemical synthesis.

Despite the problems and inherent limitations, the field of biotechnology will greatly increase the possible compounds available for drug development and, in fact, promote

development in many novel areas that have lacked effective therapies. The initial cost and the pressures for successful development are so great that the critical elements of the development process will need to be more effectively used if we are not to repeat the mistakes of yesterday. For example, demonstrating the blood clot lysing capabilities of rTPA and the reversal of an acute MI *in development* were not enough for commercial viability of the product. Genentech persisted and undertook to perform a comparative study of rTPA with streptokinase, demonstrating superiority of the rTPA product. The superiority of the rTPA combined with an aggressive marketing strategy permitted Genentech to dominate the thrombolytic market. In the development process of biotechnology products, their value and place in the therapeutic armamentarium may be as important as the demonstration of efficacy in a pivotal trial. Discussion of the pharmacoeconomic impact of new therapies must take into consideration the great expense of a biotechnology-derived drug, the benefits of the drug, its place in therapy, and especially its cost-benefit ratio as critical factors in the product's success.

Recently, the area of oncology has been most promising for success in translating biotechnology into successful products. A number of specific products targeted to cancer have proven successful in clinical trials, although possibly limited in overall impact on the treatment of cancer. Specifically, antibodies used to target specific receptors, turning off cell growth, appears a fruitful approach in the oncology area. Still the potential for proteins to affect genes and modify disease is tremendous, and we are just touching the surface of the myriad of potential opportunities this field has to offer. Especially interesting, but disappointing to date, is the field of angiogenesis, which involves inhibiting angiogenesis and tumor growth or increasing angiogenesis and, thus, decreasing organ ischemia, such as that seen in coronary artery disease. Perhaps these areas will become fruitful in the coming years.

31.5 Gene Therapy

Gene modification, substitution, and inhibition are a promising array of new strategies for the effective treatment of disease. That a single gene may be responsible for a metabolic disease, like gout or homocystinuria, for example, seems reasonable. That a single gene mutation could cause a condition like Ehlers–Danos syndrome also seems reasonable. But that breast cancer, lupus erythematosus, or coronary artery disease might be caused by a single, abnormal gene is surprising, to say the least. Yet, a mounting body of evidence supports these claims. These developments are exciting and may represent a new age of effective therapies for some of the most difficult conditions to treat in humans. But the identification of the gene itself, although an important first step, is only the initial movement of a long process toward the cure of a disease. The techniques for gene modification are rudimentary and certainly need further study.

An area of cardiology in which gene therapy should be most promising is restenosis following acute angioplasty. Angioplasty entails placing a catheter in the coronary vessel, inflating a balloon at the tip of the catheter and pushing aside the atherosclerotic lesion. This is rather a successful technique; however, a major problem limiting the success of angioplasty is restenosis. At the time of the initial angioplasty, there are stimuli that initiate cell proliferation of the media, which leads to restenosis. The medial cells that proliferate are very homogeneous, and this process seems to occur quite rapidly in about 20 to 60% of individuals having a single-vessel angioplasty. But even this simple model for gene therapy has proven a difficult target. There has been some interesting work done with antisense therapy in the

area of restenosis. Clinical trials are currently in progress using antisense material. However, so much of the methodology is new, which impedes clinical development. The use of a viral vector to insert the material in the medial cells to turn off protein synthesis is limited by concern about the use of viral vectors. Major questions arise. Can the virus replicate? Will the gene be correctly inserted, or will additional genetic material of the virus be inserted? Thus, validation and safety aspects are formidable and can markedly slow the development process. As experience increases with product development and manufacturer skill in conducting clinical trials, the overall time for developing gene-manipulation strategies will decrease. Recently, a report on the effectiveness of probucol in preventing restenosis is an example of the role serendipitous observations still play in the development of new therapeutic approaches.[3] The changing therapeutic landscape also influences commercial development. So whereas antisense therapies and other agents slowly develop, the advances with drug-coated stents may render new therapies for restenosis moot. We are only at the initial frontiers of gene manipulation. The possibilities are phenomenal. Whether the promise is realized cannot be answered at this time, but the concept of preventing or even treating serious diseases like cancer in advanced stages or incurable conditions is so exciting that it makes the concept scientifically irresistible.

A worrisome problem that has developed is the penchant to patent a gene even when our knowledge of what it does is only partial. Supporters of gene patenting believe that the proprietary nature of a patent will spur commercial development. However, a gene is not a drug but part of the disease process. We do not patent diseases. One therapy directed at a given gene may be different from another. An antibody could block gene expression, a chemical could block a protein from being made by a gene, or a gene's action could be promoted instead of inhibited. But if the gene is patented, then many facets of research can be blocked by the patent owners, blocked not just by the cost of working with the gene from royalty payments, but deliberately, by the gene's owner, who wants to ensure that his or her therapy directed at the gene is the only one available. Patenting genes will, I believe, be a major inhibitor of research in this area and a grave error on the part of the U.S. Patent and Trademark Offices and the courts, which have so far upheld this concept of patenting genes.

31.6 The Developers

Along with the evolution of the discovery process and the tremendous influence the revolution in information handling will have on drug development, changes in the participants in the development process will also have a considerable effect. There is an evolution occurring in the parties to the development process. Observing the trends makes me think of theories about the origins of the universe with oscillations in mass accumulating, exploding, and reaccumulating, forming large aggregates and small breakoff components. The drug discovery process started, perhaps, with the entrepreneurs who led the field successfully and developed the large corporate giants of today. Merck & Company (Whitehouse Station, NJ, U.S.A.), Hoffmann-La Roche (Basel, Switzerland), and Pfizer (New York, NY, U.S.A.) are examples of one-man entrepreneur-driven companies expanding into major, international concerns. In fact, the major companies dominate the pharmaceutical industry to an unparalleled extent.

The largest ten companies represent more than 90% of pharmaceutical sales. During the last 20 years, mergers have continued, and the last few years have seen even further consolidation of the pharmaceutical industry. One analyst reported in the *Wall Street*

Journal that for a company to survive, it must be able to compete with the major players in the pharmaceutical field.[4] This is not just because of funding requirements for drug development programs, but because of the development impediments established by these very large competitors. Impediments can be something simple, like the number of patients exposed to a new entity, or more complex, such as a survival study or the use of experimental ancillary technologies that are prohibitively expensive and would not be automatically required for the development of a compound. These impediments can create an impression, both to the FDA and within other companies, that they are requirements for developing a second or third agent in the field, making such development much more difficult, time consuming, and expensive. The time factor is especially important because the longer it takes to develop a compound, the more dominance in the market the first drug has gained. Time is the same as money, and the loss of a product lead can all but destroy the market potential for an agent.

Another result of industry consolidation is the tremendous pressure to develop "blockbuster" products. To sustain the pharmaceutical behemoths, one billion dollar compounds are mandatory. Compounds that only gross 50 to 100 million are no longer of interest to the large companies, and the "niche" diseases they treat are no longer attractive targets for therapy. A company that loses one or two compounds in its pipeline can rapidly fall from favor and become a possible takeover target. The recent failure of Merck's pipeline has led to a crisis at that company, a crisis that questions the very survival of the company. However, the focus on "blockbuster" products has created an opportunity for small companies to develop drugs in the vacuum created by large multinational companies. An opportunity not just to develop "niche" compounds but also "blockbuster" therapies. An insightful discovery is not limited to large robotic labs. "Big science" is not the substitute for new ideas, new approaches, and "thinking out of the box," which often are the requisites for new discoveries and blockbuster products.

31.7 Small Development Companies

Considering these obstacles to development and the considerable regulatory maze, the trend to conglomerate with bigger and bigger companies is not surprising. What is indeed surprising is the simultaneous opposing trend of the development of the very small niche start-up companies proliferating along with the ever-increasing size of the major players in pharmaceutical development. In fact, it is not just how small these companies are that is important; it is that they only encompass an aspect of the drug discovery and development process. Some companies focus on discovery; others specialize in clinical development.

Some companies plan to license their product to a larger firm for marketing once the new drug application (NDA) is granted. Then there are other companies that neither discover nor develop drugs, but rather market developed pharmaceuticals. One may argue that the niche enterprises are doomed to failure. But a number of factors combine to make this approach viable. The total overhead of the very large companies limits development to compounds expected to have sales of at least 100 million dollars. At times, a drug will have a smaller market, and a large company will develop the product for public relations, because of interest in the field, to bolster sales of existing products in their product line, or out of sheer miscalculation of market potential. However, the small companies look at a potential market of 5 to 50 million dollars a year as a bonanza. Their costs are far lower, permitting the recoup of adequate profit margins after developing, marketing, and discovery costs are accounted for, and expenses for sales are covered. Small companies cannot carry

out clinical trials at the same level expected of a Pfizer or a Merck. Thus, the studies are fewer, smaller, and aimed at proving the efficacy and safety as directly as possible. Clearly, however, the niche company can play a major role in drug development. They service areas not considered appropriate in terms of market size by the larger companies. They represent the dynamic growth of academic entrepreneurs who look to the commercialization of their ideas, especially in biotechnology and gene manipulation.

31.8 Academicians and Entrepreneurs

Many companies are investor-driven and, thus, have an intense dedication to development and success. But can a company with one or two products compete? It appears it can. However, often the process requires merger or acquisition by larger firms to succeed. Some companies start out as "spin offs" of academic research, others license a product being developed by "big pharma," and others start as marketing or generic firms that grow from their own success.

King Pharmaceuticals (Bristol, TN, U.S.A.) started small, made some brilliant acquisitions, and through mergers, has moved into research and development. Biovail (Mississauga, Canada) is an example of a generic company now developing its own products and becoming a small, viable concern. In the biotechnology field, most companies are small. The biotechnology industry seems most appropriate for the small-company approach. Whether companies besides Genentech and Amgen can climb into the "big" pharma group remains to be seen. Centocor (Malvern, PA, U.S.A.), Genzyme, and U.S. Bioscience (Philadelphia, PA, U.S.A.) have all made attempts to move into the second tear of "pharma," some more successful than others, but it is difficult to develop a viable product and then to sustain research and development and continue to grow. But the trend is clearly established.

Academicians with a novel idea no longer publish their results and go on to the next project. Rather, patents are obtained, and a company is started. I marvel at the recent reports of a new technique in cardiothoracic surgery being performed using the laparoscopic approach. Instead of reading about the advances in the *Journal of the American Medical Association* (*JAMA*), the discussion occurs in a recent issue of the *Wall Street Journal*, focusing on the possible initial public offering (IPO) that will be forthcoming for a company making the instruments essential for the procedure.

Entrepreneurs and academicians are forming alliances that may speed a procedure or chemical entity into a viable product for development. Although the free exchange of ideas may be limited, and scientific discourse suffers a bit, the possibility of widespread clinical use facilitated by commercial development is enhanced. The pros and cons of this approach are not for us to debate, but rather to accept as a trend that is ongoing and growing. I do think that the fast-moving nimbleness of these small, dynamic companies, coupled with their lower overhead costs, offers considerable benefit to pharmaceutical discovery and development. Drugs are being developed that the larger concerns would not have considered. The advancement in niche areas, such as orphan drugs, is, for the most part being pursued by smaller companies. I believe this is a healthy trend and one that will force the industry to streamline.

Combined with the trend of small niche companies in drug development has come the parallel corporate trend of downsizing and the hesitancy to expand divisions to take on temporary projects. A great number of large companies are contracting for critical aspects of drug discovery and development. Compounds can be manufactured under contract.

Consultants can put together manufacturing specifications, and preclinical testing and stability work can be done under contract. Clinical studies are performed by clinical research organizations (CROs) with the data handled by contract statistical analysis. A consulting team can put together a NDA under the supervision of a small, core group at corporate headquarters. This can be done for the small company or the very largest of the pharmaceutical giants. Parts of a project can be subcontracted. Indeed, it is not uncommon for intermediate-to-small projects at the largest companies to be entirely subcontracted. For these reasons, the CROs and other contract service organizations (CSOs) have become successful. A bonanza of new business has created exponential growth for these types of companies. The companies are competitive, and the work is relentless, but results are what make the industry thrive, and drug development has, in some instances, sped up using this "piecework" approach. To some extent, the "virtual" pharmaceutical company has materialized.

There are dangers with the fragmented approach. Outside companies can be less dependable, projects can fall apart when the capitalization of the company is inadequate and they go under, less than favorable schedules can sometimes develop because the project is not necessarily the highest priority of the contracting company. The fragmented approach can create situations in which the contracting company may be less alert to important clinical findings that should alter the development program or to serious toxicities that need to be taken into account.

If studies are performed outside the United States, as they often are, the quality of the data and the important aspects of clinical study acumen by the site investigators are often lacking. Important information about the drug may not be passed along, and this can seriously impede the development process. In addition, corporate rapport with the site investigators may be lost, and the important "seeding" of the market with experienced investigators with a product may not occur when contract organizations are involved and non-U.S. data are employed. However, there may be significant cost savings and increased patient accession with the CRO and foreign study approach that may make using them advisable.

Clearly, a balanced program that gives careful consideration to the limitations of using CROs — to run the studies, provide the statistical analyses, monitor the processes, and even, coordinate the studies performed outside the United States — needs to be evaluated against the more traditional approaches to drug development. The CRO in clinical research is discussed in detail in Chapter 21.

31.9 Government, Regulation, and Drug Development

The influences of the federal government are pervasive in our society, from our tax structure to the actions of regulatory agencies. All aspects of industry, and especially the pharmaceutical industry, are greatly influenced by the government. In the 1990s, some manipulative politicians targeted the pharmaceutical industry in their rhetoric to pander to voters. This populist approach has continued, especially from leading Democrats (but from some Republicans, as well) — all prone to this approach to garner votes. But, for the most part, there is a balanced tension between the Democrats, representing more government, and the Republicans, representing less government and increased deregulation. This is, of course, a simplification, but one with historic justification. Clearly, there is a trend against government as the provider of solutions. How the trend will develop in the short term is difficult to predict.

Even with the progression away from government and regulation, the effect this trend on the pharmaceutical industry remains substantial. The loosening of Occupational Safety

and Health Administration (OSHA) regulations and environmental impact statements are more likely than changes at the FDA, despite many discussions by the Republican Congressional leadership regarding FDA. The specter of an unprotected public, however, is a difficult "political cry" to oppose, and one, I believe, not readily challenged except by the most ardent of conservative Republicans.

The pharmaceutical industry (especially the larger companies) appears to support FDA rather vigorously. Big "pharma" operates successfully within its framework, and in a way, the FDA has become part of the process of limiting competition and diminishing the effectiveness of smaller companies unable to compete against the more-formidable pharmaceutical giants. FDA also serves the industry well, particularly in the scientific arena, ensuring efficacy and safety and instilling high, public confidence in pharmaceutical products. In addition, FDA stands at the "gate," the guardian of "big pharma's" market — both have a stake in preventing the reimportation of pharmaceuticals.

A major "zig-zag" in the political landscape has been that a Republican-controlled Congress has enacted a Prescription Drug Benefit. This legislation will neutralize the "Medicare issue" that was a major underpinning of the political attractiveness of the Democratic Party to large numbers of voters, especially elderly voters. However, Republicans have created one of the largest social entitlement programs of history, expanding not diminishing the role of government in our lives. Politics make for strange positions and alliances. Because the Prescription Drug Benefit was created by Republicans, the Bill intrudes less on the marketplace and no price controls have been imposed. Still, with the federal government being the main purchaser of pharmaceuticals, it may only be a matter of time before we see subtle, and then not so subtle, price controls to control cost. Democratic opponents are already advocating these measures. But few think of how this will affect industry and the quest for new drugs. If price controls were placed on cars, we would have fewer new models and less variety, and we would be driving vehicles analogous to those nostalgic days of the East German "mini's" or those classic boxes from the days of the Soviet Union.

The cost of pharmaceutical research is staggering, the success rate poor, which makes pharmaceutical research a very risky business. Today the U.S. free market is what drives the engine of drug discovery and development. United States, European, and Japanese companies all look to "make it" in the United States because of the price controls in Europe and Japan. European and Japanese economies are going to have to pay more for pharmaceuticals if the United States is going to pay less for new drugs. In a way this is a trade issue. Currently, U.S. consumers are subsidizing drug development. With the new, massive intrusion of the federal government into the pharmaceutical purchasing arena, there will be tremendous pressure to reduce the cost of drugs in the United States. This is acceptable as long as research continues to be funded through an incentive-based system. The U.S. government is going to face a major confrontation with Europe and Japan in this arena, which may make difference over the Iraq intervention pale in comparison.

31.10 Government Research

The importance of defense and space-related technology on drug development has been minimal and will probably continue to be disappointing. A more effective use of research funding coming from the space and defense research consortiums could be obtained by a grant system, similar to those of the National Institutes of Health (NIH) and National Science Foundation (NSF), but with more decentralization. Although the NSF and NIH are

most imperfect systems, they are far better at supporting the advancement of knowledge than the military or a space administration's approach. Whether the government turns in this direction cannot be predicted, but it may be unlikely. However, the trend toward very big scientific projects has slowed, and a more reasonable, decentralized approach is taking shape under the new Congressional leadership. This is encouraging because by supporting new programs, small programs, and diverse projects, we are more likely to see important advances, as opposed to the results seen when the established, industrialized, scientific complex and its bureaucracy are the only recipients of support.

Governmental support for pharmaceutical-related research, clinical pharmacology research, or research related to drug development is minimal. This is unfortunate because tremendous public health benefit could be obtained. This is not to suggest that government should compete with industry; but in areas that industry is ignoring or in fundamental research that leads to discovery and development or that is ancillary to drug discovery and development, government could and should play an important role. However, a major component of the nation's public health remains solely funded by for-profit pharmaceutical enterprises. The federal government's genome project appears promising for the biotechnology, gene-manipulation sector of drug discovery and development. That information is fundamental and will form the information base of many discoveries in the future. If the government patented its findings and didn't facilitate its dissemination and use of this information in research and practical product development, it would be counterproductive. That such an approach has stopped (gene patents) and the government is, once again, returning to its role as a facilitator, not a competitor, and is no longer aiming at accumulating wealth, is very good, indeed. That the government has backed away from the concept of patenting data from the genome project, in competition with the private sector, is a sign that the federal bureaucracy can be modified and be responsive to the needs of society.

31.11 The FDA Regulators

The drug development process occurs within the structure defined by the FDA. From initial clinical testing in phase 1 to later phase 2 and 3 clinical trials, the FDA has considerable influence and control, although it interferes minimally. Unlike European agencies, for example, the scientific levels of the FDA are most accessible at all stages of development from before work starts on investigational new drug (IND) applications, phase 2 trials, or NDA meetings, the FDA can provide meaningful guidance to a drug development program. FDA is the judge of the data presented and the "keeper of the regulations," but its assistance comes from experience in the drug development process. The scientific division chiefs and other senior individuals at FDA see a tremendous number of clinical trials, have encountered a variety of clinical development problems, and can without disclosing confidential information, provide considerable assistance to those involved in drug development. Whereas an individual within a company may be involved in four or five compounds during a career, the FDA senior people may see that many compounds in a week and see them from many different perspectives. Clearly, the FDA is the nexus of pharmaceutical development information and training that, unfortunately, has not been used as effectively as it should be.

Those involved in drug development must work in concert with FDA. The FDA and industry, working together on a product, will often bring about a development program that is more effective and more efficient in time and resources. Too many companies take

FDA's advice as dictum. There are regulatory advisors, what could be termed "the shadow FDA," who tell industry what FDA requires and wants, and whose recommendations are all too often distortions and impediments to effective drug development. FDA should be relied on as an important resource, one that those who are pivotally involved in drug development should communicate with directly. Regulatory advisors, consultants, past regulators, facilitators, and legal advisors all have their place, but they should not be interposed between those at the companies who are the critical links in drug development and the agency. No advice should be binding, everything needs to be discussed, and reasonable approaches need to be taken.

The individuals at FDA are not omniscient. A development plan may not work and may need substantial modification. Failure to realize this and blindly going forth after an FDA conference can lead to failure. Coming back to FDA and saying, "But this is what we were instructed to do" is foolish and, in a sense, undercuts the free and open exchange of ideas between the regulators and the developers. Advice is what is given, and later reproach because of changing circumstances, developments in the field, or just the lack of efficacy of a compound is counterproductive. In fact, it may deter the critical assistance from FDA that can be so very helpful to a drug development team.

These impressions, of course, need to be modified in the context of the divisions and the individuals involved. There are differences among and between divisions and individuals at FDA, and this needs to be factored into the equation. But, clearly, the companies most successful in development have created a working relationship with the FDA and have made use of the extensive scientific experience these individuals have with the drug development process. Having been the organizer of a course on cardiovascular drug development, protocol design, and methodology for 15 years, I can attest to the unselfish assistance of many senior individuals at FDA. Their knowledge of the drug development field, their interest in successful drug development, and their desire to find the scientific truth is clear-cut. Whereas the course involves many leading academics and industry physicians who have considerable knowledge, each year the symposium demonstrates that senior FDA participants consistently offer a broad knowledge of the field of drug development.

FDA can facilitate drug development further than it does currently. There are times that the delays are needless, that the debate is unhelpful, but the era of the "drug lag" behind Europe that so severely crippled therapeutics in the 1950s, 1960s, and the 1970s no longer exists. However, excessive drug regulation is not the goal. Rather, more expeditious, less-costly drug development should be the goal of the FDA in the information age. A case-by-case review is no longer necessary. Separately chronicling each data point and a meticulous review for efficacy and toxicity by a junior official is an immense waste of time. Having the primary reviewer recreate the NDA piece by piece and then producing his or her own summary is a laborious process that obviously can take a year or more. The NDA is put together by hundreds of individuals highly trained in the pharmaceutical industry, and having one or two people go through this on a line-by-line basis, checking every data point, reanalyzing the presentations, is going to be a most arduous and time-consuming process.

Quality assurance techniques are in place to ensure the accuracy and integrity of an NDA database. The FDA could make use of these techniques, and clearly, it will need to strengthen procedures by the end of the century, applying sophisticated computer techniques to make analyses as expeditious as possible. To keep up with the information age, the FDA will be one of the links in the drug development process that is most stressed by forthcoming change. User fees and more FDA revenues are not the answer; placing the cost of submission beyond the capacity of small startup companies is ill advised. Using these funds for more and more reviewers, thus expanding the laborious approach to data

review is fallacious. Many of these programs are not objected to by the giants of industry and, in fact, are encouraged to the place impediments on the more formative, dynamic, small companies. This is another of the anticompetitive practices in which the FDA has become an unwitting ally.

The statement that the FDA needs to "take its time," to "plow through each data point," to protect the public is one that is heard often. By never approving a drug, FDA would be the most protective, because no adversity would ensue from approved drugs. The adverse effects that would result from having no therapeutic advances would be intolerable. Thus, a compromise in the tension — between the regulators charged to protect the public and the public's need for new, effective therapies — needs to be struck. The use of the information revolution to facilitate drug development needs to be explored. We are at the beginning of this exciting period, and the government will evolve more slowly, perhaps, than other centers in the development process, but it will evolve.

A number of approaches are possible. The use of quality assurance techniques for partial-to-comprehensive data verification on a random basis certainly needs to be validated and then applied. Perhaps data analysis performed by certified groups that are paid for by the company but, at the same time, are licensed by the FDA, would eliminate the need for data reentry and reapplication of analytical techniques. Focusing on quick review techniques for the critical, pivotal studies and ascertaining veracity needs to be placed at the top of the review list. With the acceptance of efficacy, rapid computerized analysis of the product's toxicity and comparison of the results to those obtained with other agents could permit an estimation of the agent's potential benefits and toxicity. This could facilitate early presentation of the NDA material to an advisory committee that would be able to understand the drugs place in the therapeutic armamentarium and decide whether a more prolonged and thorough evaluation was needed or an early release could be considered. Of course, an early release might be combined with a more prolonged preliminary period, where information is collected on adverse experiences and efficacy, and those items could then used for continued drug evaluation.

The process of approving a drug, getting very little additional information after the approval, and allowing the drug to remain on the market forever is as wrong as a very slow and time-consuming initial development process. Because it is so difficult to get a drug off the market and we have so little postmarketing information reinforces the regulators' need to make the initial approval so stringent. It would be far better to look to a system, such as that in Great Britain, in which there is a provisional stage of approval, with a detailed program of postmarketing surveillance that encourages participation by practitioners. We must understand that drugs can be marketed, and then knowledge and information may develop that changes our initial impression. A drug could be severely limited in its labeling, have warnings issued to physicians who will be using it, or even taken off the market if our knowledge about the product changes. A drug withdrawal should not be looked at as a criticism of the FDA, but as a realization that our knowledge continues to grow.

Unfortunately, transcripts exist of Congressional committees led by inquisitors who severely criticize regulators when adversity is later discovered from an agent that was approved on quite a meritorious application. If we can get the ground rules straight and be able to understand that our knowledge base expands constantly, we will understand that difficult regulatory decisions may be necessary with new knowledge. With this understanding we can accept the early approval and later withdrawal of a drug, without faulting our regulatory colleagues. This approach is necessary if we are to fundamentally change the drug development process for the better. This is a difficult change, because so much of the process is developed by lawyers who view drug development with an eye to

litigation, looking for guilt or innocence, rather than a pursuit of science, where knowledge is continuously changing as more information is collected, and there is no guilt or innocence, no liability to determine.

31.12 Conclusion

Drug development and discovery is a most exciting field. It is creative, intellectually taxing, and organizationally demanding. Those involved are to be congratulated for undertaking efforts that are usually anonymous but that affect clinical therapeutics to a considerable degree. The drug discovery process has gone through the botanical phase and the synthetic chemistry phase and is now, into a most exciting era of biotechnology and gene manipulation. There is an awful lot of hoopla in this arena, but we must remember, it is just one of the rings of a three-ring circus, that is part of the political process.

Our understanding of medicine, disease processes, and the statistical evaluation of clinical trials has evolved tremendously. These are ongoing arenas that are important and that dynamically interact with the drug discovery process. Another arena, the drug development program itself, is being markedly affected by the third wave, the information age. We can only think back to when a chapter like this would be handwritten, typed on a typewriter with carbon paper, corrections made, and a copy sent off to the publisher. Word processing has revolutionized this approach and will continue to revolutionize it in the next couple of years. This same revolution totally changes the drug development process, facilitates it, and further facilitates the review of the database presented. All phases of drug development will undergo radical change, and we will be better for it. This is a most exciting era and one in which participation will be both enjoyable and rewarding.

References

1. Toffler, A., *Future Shock,* Random House, New York, 1970.
2. Tardif, J., Probucol and multivitamins in the prevention of restenosis after coronary angioplasty, *N. Engl. J. Med.*, 337, 365–372, 1997.
3. The SOLVD Investigators, Effect of enalapril on survival in patients with reduced left ventricular ejection fractions and congestive heart failure, *N. Engl. J. Med.*, 325, 293–302, 1991.
4. Tanouye, E., Does corporate funding influence research?, *Wall Street Journal*, B1, B, January 8, 1998.

Section VI

Case Histories

32

The Discovery of Rituxan

Mitchell E. Reff

CONTENTS

32.1 Introduction

The therapeutic promise of monoclonal antibodies had not yet been achieved when I arrived at IDEC Pharmaceuticals (now Biogen Idec, Inc., Cambridge, MA, U.S.A.) in May 1990. Monoclonal antibodies had been identified as potential magic bullets when they were first identified in a seminal *Nature* publication in 1975, which suggested that their high antigen specificity and affinity would make them very suitable for human drug development.[1] Within 9 years, the Nobel Prize in medicine was awarded to Drs. Kohler and Milstein for their discovery. Still, by 1990, 15 years after their initial work, only a single monoclonal antibody, OKT3, had been approved for a therapeutic application in the United States. OKT3 was used primarily in kidney transplantation to help prevent transplant rejection. It would be 1997 before the next intact monoclonal antibody would be approved for a therapeutic application in the United States. Approved to treat non-Hodgkin's lymphoma (NHL), that antibody was Rituxan® (rituximab), the first monoclonal antibody approved in the United States for the treatment of cancer. Until the 1990s, when a series of advances

opened the door for their development, persistent scientific, technological, and economic barriers stalled the therapeutic potential that monoclonal antibodies seemed to offer. Rituximab was just the first in a wave of what is now more than 400 monoclonal antibodies in clinical trials for a veritable host of indications. This is its story and the story of the team at IDEC that discovered Rituxan and developed it into an effective therapy for NHL.

In 1990, IDEC was a company of only 60 people, but it had already begun conducting clinical trials. IDEC which had been manufacturing monoclonal antibodies from hybridomas under good manufacturing processes (GMP) and purifying the monoclonal antibodies under GMP, had received approval from the FDA to test the monoclonal antibodies in humans and had initiated clinical trials. Rather than a traditional biotech company consisting only of research scientists, IDEC was already a pharmaceutical company in miniature, with clinical, manufacturing, quality, regulatory, and administrative staff. The entire research and preclinical organization comprised less than 30 people at this time. To make monoclonal antibodies an economically viable therapy, the team at IDEC would have to express antibodies at a higher rate than had been previously achieved. They would also have to maintain the safety and efficacy of the antibody when produced by large-scale manufacturing and conduct clinical trials for a therapy that had no established precedent. In short, much of IDEC's future was staked on a vision of drug development contained within the high-intensity focus of a small biotech firm.

My scientific background, before IDEC, had primarily been in high-level expression of recombinant proteins in mammalian cells, but I had not yet expressed any antibodies by that method. Nevertheless, when I was interviewed for the position at IDEC and was questioned about the possibility of making antibodies as recombinant proteins in mammalian cells (rather than in hybridomas, which are fusions between a B cell making a particular antibody and a tumor cell line), I remember optimistically stating that it would be no problem to make "boatloads" of antibodies. As it turned out, we were able to do so, and achieving high-level expression of antibodies enabled the administration of grams of recombinant proteins to humans in a cost-effective manner, one of the reasons for the therapeutic success of rituximab. The adaptation of existing scientific techniques to bridge traditional barriers in monoclonal antibody drug development was characteristic of the rituximab development process. Early anticipation of scientific and regulatory needs, combined with staff and corporate flexibility, helped IDEC develop a novel therapy; and Rituxan, in turn, shaped IDEC. In 1995, when IDEC signed with Genentech (South San Francisco, CA, U.S.A.) to finish the development of rituximab, it had less than 200 employees. By 2003, just before its merger with Biogen, IDEC had 1378 employees, of whom 129 comprised the research and preclinical organization.

32.1.1 Monoclonal and Chimeric Antibodies

The typical antibody or immunoglobulin is a protein, normally found in the blood, which consists of two large, identical polypeptides (called the heavy chains) and two smaller, identical polypeptides (called the light chains). All four chains are connected by disulfide bridges, the heavy chains to each other, and each light chain to a single heavy chain (Figure 32.1). There is a *variable region* of the antibody in both the heavy and light chains. The variable region is a part of each antibody that creates its unique binding specificity. This binding specificity is the property of the antibody that allows it to recognize and bind to specific pathogens such as bacteria, virally infected cells, or tumor cells. Normally, an animal's antibodies do not recognize its own healthy tissue; when antibodies bind to normal cells in humans they can contribute to autoimmune disorders such as rheumatoid arthritis and systemic lupus erythematosus (Lupus).

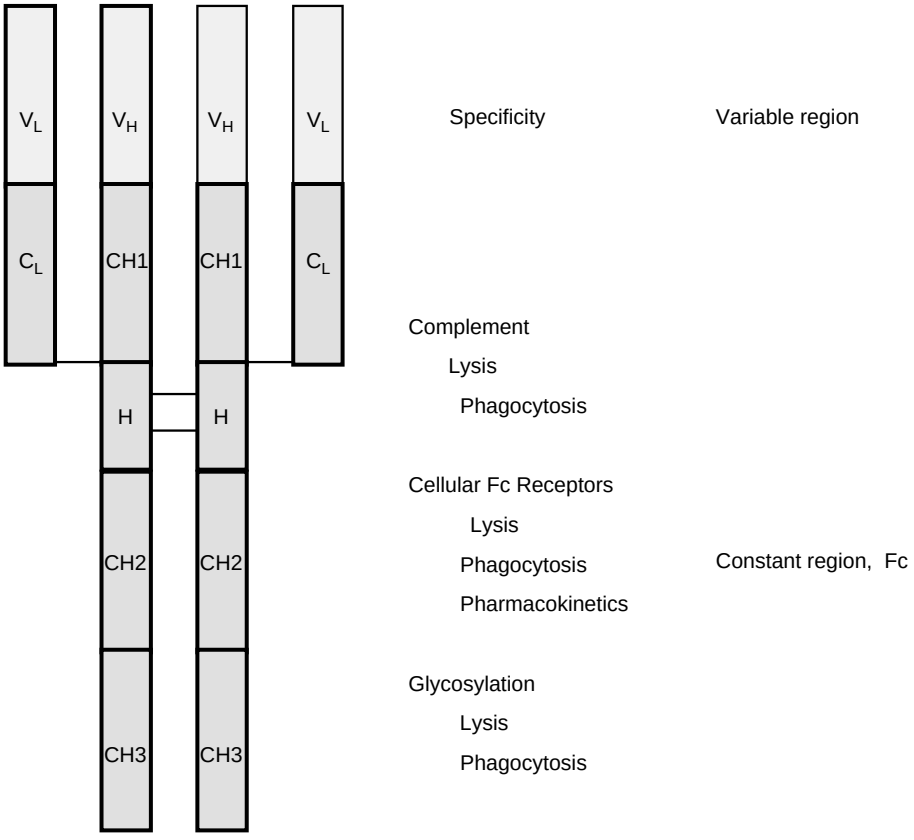


FIGURE 32.1
Antibody stick figure.

The other part of the antibody molecule, called the constant region, can interact with multiple proteins in the animal. These include interactions with a serum protein (C1Q, the first component of a set of interactive proteins in the blood named complement proteins) that cause a series of additional complement proteins to bind and form a large protein complex bound to the antibody and the pathogen. This complement interaction can lead to a pathogen being engulfed (phagocytosis) by cells that have complement protein receptors on the surface. Two different types of white blood cells — neutrophils and macrophages — can perform complement-mediated phagocytosis in this manner. Alternatively, the complement cascade can form a protein complex on the surface of a bound cell and poke holes in the cell leading to cell death. This form of cell death is called complement-dependent cytotoxicity (CDC).

Other receptors, known as *Fc receptors* — present on white blood cells (natural killer cells, neutrophils, and macrophages) — can directly recognize and bind to the constant regions of an antibody (the Fc region) when it is bound to a pathogen. This recognition of an antibody bound to a cell can lead to that cell's death by lysis (rupture of the external membrane by natural killer cells) or by phagocytosis. The lysis of a cell caused by an antibody binding to it and then interacting with a second cell with Fc receptors is called antibody-dependent cytotoxicity (ADCC).

Antibodies also have a very long circulating half-life in the blood when compared with other proteins. There is a receptor (*FcRN*) on endothelial cells lining the arteries and veins

responsible for this half-life, which can be as long as a month in humans. The long circulating half-life, the ability to attract other defense mechanisms, and the binding specificity of antibodies make antibodies particularly suitable for human drug development. But the polyclonal nature of immune response to any antigen and the difficulty of purifying or manufacturing a single antibody stood in the way of developing antibodies for human therapy.

Humans have millions of different antibodies with different variable regions circulating in their blood. Each particular antibody is made by a single cell (or small number of closely related cells) called a B cell. Many early studies on antibodies were performed with the blood of individuals who had a tumor (multiple myeloma) derived from a single B cell making a single antibody. The Nobel Prize in Medicine for 1984 was awarded for the Kohler and Milstein's groundbreaking work in 1975, in which they fused a B cell — which has a limited life span and produces a single antibody — with an immortal, tumor cell line. This fusion formed a hybridoma, a new immortal cell line that produced a single antibody. The original hybridomas secreted the first true monoclonal antibodies (truly monoclonal in that the antibodies were derived from the identical offspring of a single, cloned, antibody-producing cell). Monoclonal antibodies have specificity to a particular antigen, are identical because they are synthesized by the same cell line, and can be produced indefinitely because of the immortal nature of the hybridoma. These properties make monoclonal antibodies intriguing drug candidates for developing human therapies.

The first monoclonals were murine (mouse) antibodies, because the early experiments involved injecting mice with an antigen, harvesting their B cells, and then fusing a harvested B cell with an immortal tumor cell line. There were immediate attempts to make therapeutic drugs from murine monoclonal antibodies. Scientists at Harvard Medical School tried to apply murine monoclonal antibodies for treatment of tumors as early as 1980.[2] But with a very few exceptions, murine antibodies failed as drugs in humans.

There were several reasons for these failures. First, the human body recognized murine antibodies as foreign. Human-administered, murine monoclonal antibodies made antibodies to the murine antibodies (human anti-mouse antibodies, HAMAs). This response led to inactivation of the monoclonal antibody when it was administered to the patient more then once or twice. Second, the half-life of a mouse antibody in humans is much shorter than the half-life of a human antibody in humans, usually lasting only a few days rather than weeks. This means, much larger amounts of mouse antibody would have to be administered to patients to reach levels where the antibody would be effective. Finally, antibodies are very large molecules, molecular weight (MW) of 100,000 or more, compared with the 300 to 500 MW range of many other chemical pharmaceuticals. Because of their high MW and the complex sugars attached to their amino acids, which are important to their function, antibodies cannot be chemically synthesized in a cost-effective manner; nor do hybridomas produce enough antibodies for cost-effective therapy. Antibodies can only be made economically by recombinant DNA technology in living, complex mammalian cells. The science involved in making large amounts of recombinant proteins from those complex cells was extremely primitive in the 1980s. (Nor was it a solution to pursue antibodies in human hybridomas. It is very difficult to construct human hybridomas, both because of the obvious ethical problem involved in injecting humans with antigens and harvesting their B cells and because of the scientific problem of developing an auto-antigen response. This has now been circumvented, somewhat, by mice with human-antibody genes and by bacterial viruses that present the variable regions of human antibodies. In any case, the limited protein yield of hybridomas still poses a problem for human drug development.)

Consequently, two branches of scientific research — the recombinant DNA technology required to make the mouse antibodies more like human antibodies, and the high-level expression of recombinant proteins in complex mammalian cells — had to come together to make antibodies into successful therapeutic drugs. Of course, another drawback of antibody drug therapy is that the antibodies cannot be taken orally because they will be digested like any other protein. This means that antibody drug therapy could not initially be self-administered. At first, antibodies were administered by sterile infusion directly into the blood. Today, subcutaneous administration of large proteins is becoming more common, enabling the potential for self-administration.

To get around some of the problems associated with the therapeutic application of murine monoclonal antibodies, scientists envisaged making the mouse antibodies more humanlike. Several groups of researchers genetically combined the murine variable regions of the antibody (the regions that provide antigen specificity) with human constant regions (those that allow complementary immunological interaction).[3] These combinations were called chimeric antibodies (Figure 32.2), named after "Chimera," a monster from Greek mythology made up of parts of many different animals. Chimeric and humanized antibodies maintained the specificity of the murine antibody, but now possessed human-effector functions because the constant region of the antibody was human. These antibodies interacted with human Fc receptors, human complement protein C1Q, and the human FcRN receptor. Today, humanized (a less mouse-like, more human subset of the initial chimeric antibodies; see Figure 32.2) and all-human antibodies have been approved for use in humans.

Chimeric antibodies can be expressed in mammalian cells by introducing the recombinant DNA that codes for the antibody into the cells and selecting for the rare cell that has integrated the DNA into its own chromosomes and then expresses a new protein — the desired antibody. Additional genetic manipulations of the mammalian cells are then required to select cells that secrete large amounts of the recombinant protein. Following this research and development model would theoretically lead to successful monoclonal therapies, but in practice, it had never been accomplished before the IDEC team developed rituximab.

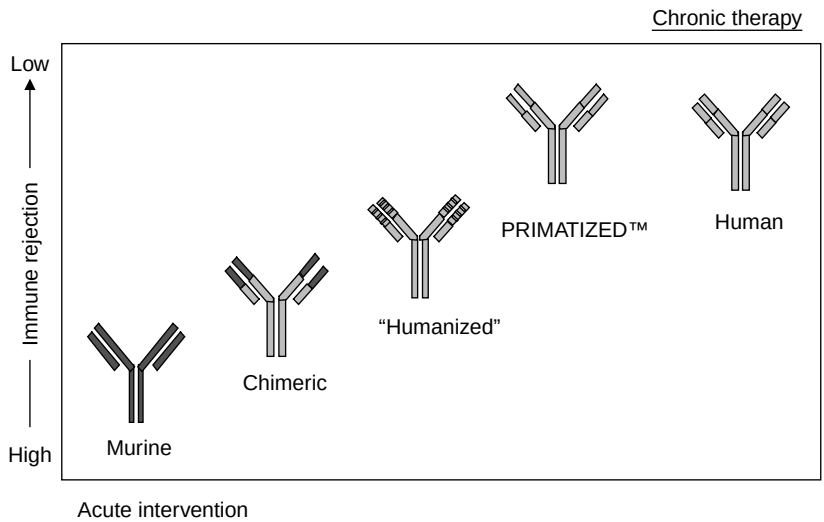


FIGURE 32.2
Generations of antibody technology.

32.1.2 Starting Down the Path to Rituxan: High-Level Expression of Recombinant Proteins

In 1990, a small group of individuals at IDEC began building the DNA molecules (antibody-expression vectors), which, when introduced (transfected) into mammalian cells, would express antibodies. There were 19 individual cloning steps between the original expression vectors, developed at Smith Kline & French Laboratories (Philadelphia, PA, U.S.A.) and then licensed to IDEC, and the modular antibody expression vectors used to make rituximab. These antibody expression vectors are modular in that they can be adapted to combine any nonhuman or human-isolated variable domains to express chimeric, humanized, or human antibodies in a given mammalian cell line. This modularity allowed the expression vectors to be used for any of the antibodies IDEC might be working on — as the vectors were tailored for high-level expression rather than for any particular antibody, a highly desirable flexibility.

The first two antibodies our group tried to express in this vector system within mammalian cells had mistakes in the genetic material that coded for variable regions and, consequently, did not express any antibody at all. The first antibody that was produced by recombinant technology at IDEC using this vector system was in early 1991. It was supposed to be a chimeric anti-CD4 antibody, but the antibody secreted by the cells did not bind to CD4. CD4 is a surface molecule on the leukocytes known as T cells. A primatized antibody (a chimeric antibody where the variable domains are isolated from a cynomolgus monkey; see Figure 32.2) to CD4 was produced later that year.[4] That antibody ended up in clinical trials in humans at the same time as Rituxan.

Two different mammalian cells were chosen to express the monoclonal antibodies: a mouse myeloma cell derived from a tumor of a mouse B cell (SP2/0) in which many academic labs had expressed antibodies, and a hamster cell in which many other recombinant proteins had been expressed (Chinese hamster ovary, CHO), but in which only low-level production of antibody had been reported when I arrived at IDEC.[5–8] Attesting to the rapidity of the scientific work being done in gene expression, that observation was overturned scant months later. High-level production of monoclonal antibody in CHO cells produced by genomic amplification using a linked dihydrofolate reductase (DHFR) gene and selection with its competitive inhibitor methotrexate (MTX) was reported in November 1990.[9] This observation proved to be significant in our later choice of cell line and method of amplification.

The scientific discussions about which cell to use to manufacture the anti-CD20 antibody that would become rituximab were intense debates. Other recombinant proteins were being made at large scale in the hamster cells, but all the academic scientists were using either mouse or rat cell lines derived from B-cell tumors (myelomas). Today we know that there are small differences in the sugar composition of the N-linked sugars attached to recombinant proteins when they are produced in different mammalian cells. Because of this, the variations between possible production cells for each recombinant protein can affect the pharmacological characteristics of the protein being expressed, and so the choice of cell has implications in drug development that extend far beyond the quantity of protein that a given cell is able to manufacture.

For example, the mouse B-cell myeloma, Sp2/0, produces a small percentage of antibodies that contain a galactose-α-1,3-galactose linkage that is not found in CHO and has the potential of being immunogenic in humans.[10] Neither Sp2/0 nor CHO contain a bisecting *N*-acetylglucosamine (GlcNAc) found in many human antibodies. The bisecting GlcNAc — which is not present on Sp2/0 or CHO cells — leads to diminished binding to FcγRIII. (Both the Sp2/0 and CHO product would be better in a cell that added the bisecting GlcNAc to the complex sugar structure.[11]) However, none of this information was available when the decision to use CHO was made. The chimeric mouse–human antibody

to CD20 (which would become rituximab) was produced at IDEC in both Sp2/0 and CHO cells and then compared for affinity (the same), ADCC (the same), CDC (the same), and finally, *in vivo* for depletion of normal B cells and pharmacokinetics in cynomolgus monkeys (the same) before the CHO cell line was chosen in May 1992 because of enhanced productivity.[12,13] IDEC's chimeric antibody to CD20 was not even the first chimeric antibody reported to bind to CD20.[14] It is still not clear why scientists at XOMA (Berkeley, CA, U.S.A.), who reported in 1987 a chimeric antibody to CD20 with human-effector functions, did not develop their antibody, although it seems likely that problems with recombinant protein expression led to their decision to not continue their research.

By early 2005, there were 18 monoclonal antibodies approved by the FDA made in mammalian cells by recombinant DNA technology, of which eight are produced in CHO, four in NSO (rat myeloma) cells, four in Sp2/0 (mouse myeloma), and two in hybridomas that are fusions between murine tumor cells and normal B cells. As our understanding of the pattern of glycosylation of the N-linked sugars (the signature of each cell line) expands, it is expected that greater specificity for desired properties effected by the glycosylation will be incorporated into recombinant protein production.

32.2 The CHO Cell Line

The CHO cell line used to manufacture the drug rituximab was produced in three steps. First, we assembled a DNA construct containing the genes that code for the anti-CD20 antibody and for those genes that would enable us to select for and coamplify the production of anti-CD20. We integrated a single copy of the heavy- and light-chain immunoglobulin genes and two other exogenous genes into the cell line. These other genes were the DHFR and neomycin phosphotransferase (neo) genes. The initial selection of cells (to see which had been successfully transfected with the construct) used neo, which allowed for the selection of mammalian cells resistant to the antibiotic geneticin (G418).

The second and third steps both involved the genomic coamplification of the construct by selecting for DHFR in the presence of MTX. The CHO cell line that we worked with is a mutant cell line produced by Larry Chasin at Columbia University (New York, NY, U.S.A.), which has been modified so that it does not have its own DHFR gene.[15] Because this CHO line is missing its copy of DHFR, it enables the genomic coamplification of recombinant genes linked to an exogenous copy of DHFR in the presence of the competitive inhibitor MTX, a technique used in the high level of expression of protein at SmithKline. Our team at IDEC was one of the first to attempt genomic coamplification in serum-free media; we attempted a variety of concentrations of MTX, eventually settling on a 5 nM concentration for the initial amplification step. Because we had isolated and subcloned G418-resistant cells with a single site of integration, not all of the subcloned cells coamplified after selection, although in our experience between 80 and 90% of the single, initial-copy integrants did amplify.

The 5 nM MTX was selected for cells that had extra copies of the linked, exogenous DHFR gene and immunoglobulin genes. This 5-nM MTX cell line, which contains only four copies of the immunoglobulin light- and heavy-chain genes, was used to manufacture the rituximab for phase 1 single-dose clinical trials and phase 2 and 3 repeat-dosing clinical trials. From the 5 nM MTX cell line, a higher-secreting cell line was isolated by selection in 50 nM MTX. This production cell line has 8 copies of the heavy- and light-chain genes on a single chromosome.[16]

Producing the CHO cell line that manufactures rituximab was unorthodox because we only transfected a single, rather than multiple, copy of the heavy- and light-chain immunoglobulin genes into the CHO cells. Normally, teams attempting high-level expression of recombinant protein aim for multiple copy, multiple-site integration. However, using an individual copy of the genes integrated into the chromosome at a single site allowed us to establish that no changes (mutations) had occurred in the light and heavy chains of the DNA construct introduced into the cells and that all of the subsequent cell lines constructed were derived from a single CHO cell.

The theory was that the later genomic coamplification would produce a more stable cell line if performed on cells with a single-site integration rather than those derived from multiple copies; even if final protein expression levels were similar, our selected CHO line would possess between 10 and 20 copies of the DNA construct, whereas selecting and coamplifying cells with multiple initial copies would result in a cell line with an order of magnitude of more copies of the construct. This latter condition would increase the chance of unwanted mutation and instability in the cell line. Empirically, our hypothesis seems to have been born out: The CHO line used to produce rituximab and other amplified cell lines have few final copies and are highly stable, even in the absence of MTX.[16]

The final cell line used to produce rituximab also has the highest reported, specific cellular productivity (50 pg cell/day [pcd]) of any commercially approved antibody in production. Most antibodies, and indeed, most proteins, cannot be secreted by any replicating mammalian cell at more than 25 to 30 pcd under any conditions. To this date, no scientist has found the "magic" cell that secretes more protein than other replicating cells; success in high-level expression has come through genomic coamplification, rather than from finding particularly robust protein-producing cell lines. Most hybridomas only secrete 1 to 5 pcd of antibody, although selection can be used to find hybridomas that secrete 15 to 20 pcd, about half the amount of the average transfected and amplified CHO cell line. Advances in titers, from batch cultures of 100 mg/L every 10 days in the 1980s to batch cultures of 1 g/L every 10 days in the 1990s (rituximab) and 4 g/L every 10 days in the 2000s (rituximab) have come from growing cells at higher densities and keeping them from dying (going through apoptosis, or programed cell death).

These remarkable increases in productivity have driven the cost of manufacturing recombinant proteins in mammalian cells from thousands of dollars per gram to hundreds of dollars per gram, at which point most of the remaining unit cost is in purification, vialing, and quality control as opposed to fermentation. In addition, the fluid in which mammalian cells grow (the media) has transitioned from being undefined and containing bovine serum to being completely defined with no added proteins except for the growth factor recombinant insulin that is produced by bacteria, a shift necessitated by the rigorous safety demands on production of recombinant proteins for use in humans. At IDEC, this change was anticipated before it became required by regulation.

As previously stated, Dr. Larry Chasin constructed the parent hamster cell line (CHO DG44) in 1986.[15] It is an aneuploid cell line with a genetic mutation (it is missing all of its DHFR genes), which allows for genomic coamplification of recombinant genes linked to an exogenous DHFR gene in the presence of the competitive inhibitor MTX.[17] The cell line was obtained at IDEC in 1988 by the molecular biologist Roland Newman. By September 1990, we had adapted it into a serum-free media, and growth was carried out in suspension culture. This change was based on my experience at SmithKline with the identical CHO cell line.

At SmithKline, the scientists in manufacturing had already expressed a preference for having every cell line grown in *bioreactors* (fed-batch fermentation tanks) and for having every cell line grown in serum-free media. (The fermentation tanks used in the industry today are basically adapted beer vats, which is why so many of my old colleagues in

manufacturing have retired to open microbreweries.) The manufacturing scientists' elected preference was a strong indicator of the greater stability and control over the line that could be achieved. It was already clear in the 1980s that adapting a cell line from serum to serum-free media changed the properties of the recombinant proteins being secreted by the cells. Changes in N-linked glycosylation have been documented in the switch from serum-containing to serum-free media. Changes in glycosylation can lead to changes in the pharmacokinetics of molecules, or to changes in the effector functions (ADCC or CDC) of monoclonal antibodies.[18] Although growing monoclonal antibodies in serum for human use had not yet been forbidden when I came to IDEC, the writing seemed to be on the wall (or in the vat, as it were).

32.2.1 Non-Hodgkin's Lymphoma and CD20

IDEC was interested in NHL, in 1990, because NHL, specifically the low-grade or follicular type of the disease, remains an incurable tumor for the vast majority of patients. It has a large and expanding patient base. NHL is the fifth most-frequently diagnosed cancer in the United States, the sixth leading cause of cancer death, and the second fastest-growing cancer (in incident rate). NHL mainly afflicts people over 50 years of age. As with most scientific research, however, there was a fair bit of serendipity involved in our choice of target.

IDEC was founded in 1985 to develop and commercialize anti-idiotypic (anti-ID) monoclonal antibodies, a patient-specific, customized approach to the treatment of NHL. Ron Levy at Stanford and other colleagues published the original idea in a paper in 1982, which showed that treatment of B-cell lymphoma with a tumor-specific anti-ID antibody was efficacious.[19] Another company (Biotherapy Systems, Mountain View, CA, U.S.A.) with the same idea was formed at the same time. Venture capital firms suggested that the two small companies should merge, which they did, in 1986. Shortly after merging, William H. Rastetter (from Massachusetts Institute of Technology, Cambridge, MA, U.S.A.; and Genentech, South San Francisco, CA, U.S.A.) was appointed as Chief Executive Officer (CEO). In 1990, Nabil Hanna (from the National Cancer Institute and SmithKline & French Laboratories) was appointed as the head of research. Both Rastetter and Hanna believed that patient-specific therapies would never be commercially successful, even though IDEC was, at that time, in the middle of phase 2 trials with a panel of 17 anti-ID monoclonal antibodies that recognized about one third of the NHL population. Primarily because of its anti-ID antibodies, IDEC had an initial public offering (IPO) in September 1991 and became a public company with a value of slightly more than $50 million.

Rastetter and Hanna were, instead, interested in pan-B monoclonal antibodies that would recognize most B-cell tumors as well as the normal B cells. Figure 32.3 shows the B-cell antigens and their relationship to B-cell ontogeny and B-cell tumors. The CD20 antigen was chosen from a number of cell-surface antigens that are specific for B cells, for a number of reasons.[20] CD20 has a high expression level on many B-cell tumors, and a very high percentage of B-cell tumors have CD20 on their surface. CD20 expression is also homogeneous so that all of the tumor cells within a specific tumor express the molecule on their surface in a similar manner. CD20 is not present on the stem cells of human bone marrow, so it was expected that normal B cells would be able to regenerate from precursor cells in the bone marrow. In addition, CD20 was not found circulating in the serum of tumor patients, so the antibody would not be bound by circulating antigens and would only bind to the tumor cells and normal B cells. Circulating CD20 has, however, been detected in the blood of chronic lymphocytic leukemia (CLL) patients today, probably as part of a membrane fragment.[21] Finally, CD20 does not move inside the cell from the cell surface when an antibody is bound to it, so antibody-coated tumor cells should be susceptible to both ADCC and CDC. Experimentation seemed to confirm this conjecture; the chimeric antibody (human $\gamma 1\kappa$)

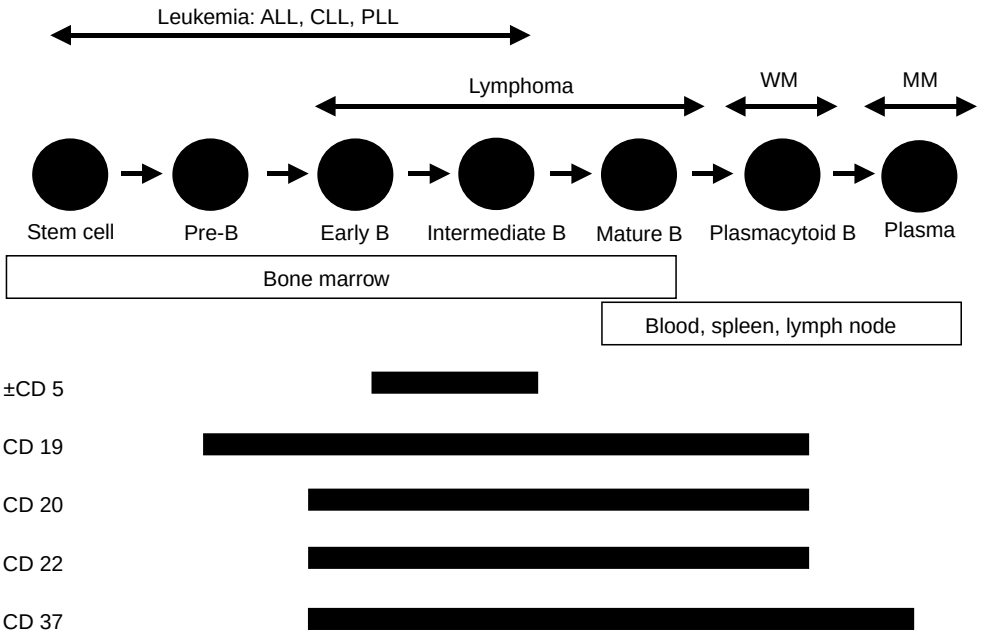


FIGURE 32.3
Antigen expression in B-cell lineage.

created at IDEC was shown to perform ADCC and CDC *in vitro*, although its murine $\gamma 1$ counterpart was inactive.[13] Experiments in the presence of human plasma, however, seemed to show transient down-modulation of CD20 by Rituxan in patients with CLL.[22]

In August 1990, Darrell Anderson at IDEC began immunizing mice with a human B-cell line, with the idea of creating a monoclonal antibody to CD20. In January 1991, Anderson identified the only hybridoma in his experiment that recognized CD20 and was a mouse monoclonal antibody of the immunoglobulin (IgG) subtype. This mouse monoclonal was the antibody from which the chimeric antibody rituximab was constructed. The variable domains of this mouse monoclonal antibody were isolated by polymerase chain reaction (PCR) by Newman's group and incorporated into the IDEC antibody-expression vectors.

CD20 is a tetraspanin molecule (one that winds back and forth through the plasma membrane four times) whose function is still not understood today. Mice that have the genes for CD20 deleted (knockout mice) have a perfectly normal immune system without any CD20 genes.[23] After IDEC had begun clinical trials, it was found that antibodies to CD20 induce tumor cells to go through apoptosis, and this is believed to contribute to Rituxan's efficacy *in vivo* (Hariharan, H., personal communication, 1994).[24] Also after clinical trials were begun, it was discovered that rituximab potentiated the effects of chemotherapy *in vitro*.[25] This observation may explain why combinations of Rituxan and chemotherapy have performed so well against cancer in humans.

Because only one of CD20's spanning loops is wholly external to the cell membrane, all the monoclonal antibodies that have an affinity for human CD20 are cross-reactive, preventing the binding of the others to the CD20 molecule after one is attached. These monoclonals must recognize the external loop differently; however, because point mutations in the loop differentially affect the binding of different monoclonals.[26] In addition, the monoclonal antibodies that bind to CD20 have different capacities to induce apoptosis, to

activate ADCC, and to cause the migration of cell-surface CD20 into a specialized portion of the cytoplasmic membrane called lipid rafts.[26] The mechanism by which antibodies to CD20 induce apoptosis remains unclear, but it does not appear to correlate with the ability of the antibody to cause migration into the lipid rafts,[27] although CDC may correlate with raft migration.[28]

Which one of these three cell-killing mechanisms — ADCC, CDC, or apoptosis — is most associated with the tumor cell death caused *in vivo* by rituximab is still unknown today. There is an observed correlation with a high-binding allotype of FcγRIII (the FcγRIII receptor is found on NK cells and macrophages) and the efficacy of rituximab.[29] This is the best data in humans suggesting that interaction of the human Fc portion of rituximab is important for efficacy. An earlier observation, which showed that rituximab had no efficacy against a human B-cell tumor in a mouse that did not contain either functional murine FcγRI or FcγRIII, also indicated the importance of rituximab interacting with Fc receptors.[30] CDC is believed to be important in treatment with rituximab because the use of antibodies to complement-resistant receptors enhances the ability of rituximab to kill certain B-cell tumor cell lines, and these complement-resistant receptors are present on NHL cells.[31,32] The presence of complement receptors on NHL tumors before treatment, however, has no correlation with response to rituximab.[33] Finally, apoptosis of tumor cells has been observed *in vivo* when Rituxan has been used to treat CLL.[34] Although in this case, the apoptosis was reported to be dependent on mitochondrial caspase-9, other studies have suggested that apoptosis of NHL lines by rituximab is independent of caspase activation.[27,34]

Whatever mechanism is functioning *in vivo* to deplete tumor cells and normal B cells, the idea that the depletion of normal B cells might have beneficial effects on treatment for autoimmune disorders such as rheumatoid arthritis (RA) was not even a fantasy in 1990. In fact, it was taken for granted that the effect of normal B-cell depletion would be negative. The question originally posed regarding a pan-B antibody such as anti-CD20 was: How much toxicity would a pan-B antibody cause by depleting all the normal B cells? We knew in 1990 that depleting all normal, helper T cells (CD4-positive), led to acquired immunodeficiency syndrome (AIDS). Would depleting all the B cells lead to infections or even the appearance of additional tumors? What would happen to circulating levels of IgG when all the B cells were depleted? Would the IgG levels decline rapidly? If they declined, how long would it take them to return to normal? How long would it take the B cells to return to normal? All these questions would have to be answered during the clinical trials.

32.2.2 The Team Drives to the Investigational New Drug

In December 1991, a project team was formed to develop the anti-CD20 antibody. Actually, two teams were formed at the same time, one to lead the development of anti-CD4 and the other to lead the development of anti-CD20. Anti-CD4, an antibody that affects helper T cells, was thought to have incredible potential to affect multiple autoimmune diseases, which were thought to be dependent on helper T cells. The autoimmune diseases included rheumatoid arthritis, and possibly, allergic asthma — both diseases with enormous markets and without effective treatment options. A nondepleting modified version of the anti-CD4 monoclonal antibody was still under consideration for additional clinical trials in 2005.

The anti-CD20 cell-surface molecule was a candidate for NHL treatment, an incurable tumor. In 1991, NHL was considered a small market, with perhaps less than $200 million in anticipated sales per year. Because the expected market size was small, there was not a lot of competition from large pharmaceutical companies. In 1995, when IDEC had impressive clinical data showing safety and efficacy with Rituxan, many large pharmaceutical

companies were approached, but they were not interested in forming a development partnership because they thought the lymphoma market was too small to make the partnership profitable. A similar pattern of large pharmaceutical company disinterest also occurred with the Amgen (Thousand Oaks, CA, U.S.A.) blockbuster drug erythropoietin alfa (EPO, Epogen). The "small" size of the lymphoma market also helped because IDEC was able to get "orphan drug" status, giving it regulatory and financial advantages in drug development.

A project team for drug development is a group of experts in diverse areas who have been charged with compiling the information necessary to file an investigational new drug (IND) application that will convince the FDA to allow testing of a novel drug in humans. Once the IND is filed, the drug development team is charged with testing the drug in humans (clinical trials) to ensure that the drug is safe and efficacious. In most large companies, professionals who have previously learned the art of drug development as a part of a project team are chosen as project leaders. At SmithKline, I had been involved as a scientist in the development of novel molecules to destroy blood clots (fibrinolytics). I had been a member of several teams developing novel drugs but had never been a project team leader. The Rituxan team was my first such opportunity.

In the early 1990s, IDEC project teams were primarily led by scientists who were pulling double duty: working in the laboratory to create and validate novel compounds and leading a group of experts in other disciplines (in large-scale manufacturing, in formulation science, in regulatory affairs, in quality control and compliance, and in conducting clinical trials). One of the reasons that development of rituximab occurred so rapidly was that all members of the project team at IDEC were the individuals who were either actively doing the work or, at most, one step removed from the science at hand. The members of the project team were not reporting secondhand or thirdhand information from within the constraints of a large organizational apparatus; they knew exactly what needed to be done, and how long it would take to do what was required. The entire team addressed the problems that arose during the development process, and individuals would often cross departments to offer suggestions and provide help, if help was necessary, for critical path decisions. The development model took advantage of IDEC's small size, using organizational cross-familiarity and cross-flexibility as powerful tools to magnify the effectiveness of the project team. The comprehensive nature of IDEC's staff allowed an extension far beyond what its size might otherwise dictate.

The time between immunizing the mice with CD20 (August 1990) and filing the IND for Rituxan (December 1992) was 2 years and 4 months. That brief span encompassed the first time IDEC had ever produced and manufactured a recombinant antibody in CHO cells for clinical trials in humans. The IND also included safety and efficacy data in primates, in which monkeys who had had their normal B cells completely depleted from their blood and depleted in part from their lymph nodes showed no adverse events.[13] In addition, a human $\gamma 4$ version of rituximab was shown not to deplete B cells in monkeys, indicating for the first time that the Fc portion of the human $\gamma 1$ antibody was contributing to depletion of cells with CD20 on their surface *in vivo*.[35]

There were several important decisions that the team made during the early development of rituximab. One was choosing the cell line CHO, rather then SP2/0, because of differences in productivity observed at the bench scale in the research laboratories. The people in charge of large-scale mammalian cell culture at IDEC — who were 400 miles away (in Menlo Park, CA, U.S.A.) from the research team in San Diego, CA, U.S.A. — had never manufactured a product from the CHO cell line. They had been growing murine hybridomas in hollow-fiber bioreactors to produce anti-ID antibodies for clinical trials in lymphoma. They had to shift not only the cell lines that they were familiar with, but also their entire methodology of manufacturing, from hollow-fiber bioreactors to fed-batch

fermentation tanks. In addition, they had to adopt new serum-free media in which to grow the cells.

Their rapid success in switching methodology and media allowed IDEC to produce large amounts of antibody for the clinical trials. The approved dosing of Rituxan, approximately 3 g per patient, is an amount of recombinant protein that previously would have never been considered because it would have been economically unfeasible using hybridoma-produced material. Andrew Grant was the process science engineer at IDEC who was primarily responsible for the incredibly rapid and successful transformation from hybridoma to CHO cell line. More remarkably, he and his team accomplished this transformation at a time when IDEC was relocating its Mountain View operation to San Diego to unify the company in a new facility. The manufacturing group had had the foresight to build a commercial-scale, 1750 L fermentation tank in San Diego; from which, rituximab would be launched in 1997.

Even today, the manufacture of recombinant protein from cell lines is still as much an art as a science, and the large-scale manufacture of proteins in mammalian cells is a delicate discipline, which, when not managed properly, has led to the demise or delayed development of many biopharmaceutical compounds. The shift from small-scale, laboratory-synthesized proteins to large-scale, protein production and purification is a transition in which failure has often had devastating effects on development. This is because the cells, which have grown attached to cell surfaces (adherent cells) in serum in the research laboratories, change the properties of their secreted proteins when they are adapted to large-scale manufacturing. The recombinant proteins secreted by cells growing in suspension in serum- or protein-free media in large fermentation tanks are fundamentally different from their bench predecessors, as discussed above. These differences are usually tied to changes in glycosylation on the proteins that can lead to potential changes in efficacy, safety, and pharmacokinetics. These issues were generally avoided in the development of rituximab at IDEC, although there were some initial problems when we shifted the manufacturing process from IDEC to Genentech.

In 1990, all the antibodies IDEC was producing for clinical trials were being formulated in normal saline for intravenous infusion. This unbuffered formulation had a shorter shelf-life than buffered solutions. In 1992, even though rituximab had not yet demonstrated efficacy in humans, Tom Ryskamp, an IDEC quality scientist, led the charge for changing the formulation of the protein, feeling that it was necessary for the long-term success of the drug. This substitution was another example of a large contribution to the drug development effort that was enabled by the close teamwork and easy flow of decision-making characteristic of the project team.

32.2.3 The Rest of the Story: From Clinical Trials to FDA Approval

At the end of 1992, Antonio Grillo-Lopez, a hematologist–oncologist, joined IDEC as chief medical officer and joined the rituximab development team. Grillo-Lopez's skill at designing and conducting clinical trials was an important factor in the rapid development of rituximab. Once the safety of rituximab was established and the efficacious dosing of four weekly infusions of 375 mg/m^2 was determined for low-grade, follicular refractory NHL, many additional, early clinical trials were conducted.[36] These trials included combinations of Rituxan and chemotherapy,[37] Rituxan combined with α-interferon,[38] Rituxan used to treat bulky disease,[38] Rituxan retreatment (patients receiving a second course of rituximab after relapsing, following an initial clinical response to the drug),[38] and patients given eight weekly infusions of Rituxan rather than four.[38]

After Grillo-Lopez joined the rituximab team, it was unusual that I remained the project leader. When most project teams enter clinical trials, a clinician becomes project leader

in place of the research scientist who has led the project to the IND filing. To this day, I am pleased that Grillo-Lopez and the other senior executives left their crown jewel in my hands because it was an experience that I will never forget. From 1993 to 1995, the rituximab project was the heart of IDEC, aptly indicated by the IDEC decision 1993 to abandon the anti-ID monoclonals, which were, at that time, in phase III clinical trials. This was a bold business move; partly because IDEC no longer had a lead product, the market response was overwhelmingly negative, dropping the market capitalization of the company to nearly nothing**.** The Rituxan team, at one point, had more than half of the people in the entire company working on some aspect of drug development for rituximab. Having a small company focus on a single drug is one way to ensure rapid development, but this does not appear to be easily reproducible.

In early 1995, Bill Rohn, the chief operating officer at IDEC, became the project leader of rituximab when Genentech signed a collaboration agreement with IDEC. The 166-patient, single-arm, pivotal trial for Rituxan began the following month, and the last patient in this trial entered the study in March 1996. The clinical results from this trial were reported at the American Society for Hematology (ASH) meeting in December 1996. The pivotal trial results were first published in 1997[39] and have been discussed in several excellent reviews.[36,38,40]

A biologics license application (BLA) for Rituxan would be filed on February 28, 1997, 4 years after the first patient was treated with the drug. The rapid acceptance of the clinical data (approved by the U.S. Biological Response Modifiers Advisory Committee on July 25, 1997 and given FDA marketing approval on November 26, 1997) was a tribute to the extremely rational and conservative approach (involving an independent panel of lymphoma experts) that Grillo-Lopez had insisted on in evaluating clinical responses to the drug.

32.3 The Surprising Effects of Human B-Cell Depletion: Rituxan Usage Expands Dramatically

Some of the most pleasant data regarding human B-cell depletion were observed in the early clinical trials. The drug was very safe. After the first infusion, in which some minor safety issues were reported (probably related to the lysis of a large number of cells in the blood), subsequent infusions showed no toxicity in most of the patients. Although circulating B cells decrease to zero following the first infusion of rituximab, there is no observed depletion of other important white blood cells, including T cells, natural killer cells, and neutrophils; there is also no observed depletion of platelets in the blood.[38] Most importantly, serum immunoglobulin levels did not decline significantly. Circulating B cells usually returned to normal levels within 9 to 12 months after treatment.[38] The reason that serum immunoglobulin did not decline significantly when B cells were depleted from the circulation is unknown. Rituximab does not deplete fully differentiated plasma cells because those cells have no CD20 on their surface. The life span of plasma cells in humans is not known, but in 1997, it was reported that the mouse has plasma cells in its bone marrow that live for the entire life span of the mouse.[41] In addition, it is probable that as serum immunoglobulin levels begin to drop, the FcRN receptor, which is responsible for maintaining levels of circulating immunoglobulin, responds by increasing the half-lives of the remaining antibodies in the blood. Most importantly, patients who have their circulating B cells depleted by rituximab do not have a higher incidence of infection or any apparent increase in other tumors. In addition, relapses due to the appearance of tumor cells that are negative for CD20 are very unusual.[42]

This remarkable and unexpected safety profile coupled with efficacy in treating refractory low-grade follicular NHL has led to the increased usage of rituximab for other types of NHL, other B-cell tumors with CD20 on their surface, and other indications where B-cell depletion may have a desirable outcome. As a result, Rituxan reached yearly sales of more than a billion dollars in 2002, and its sales are still expanding. One example of the expanded use of Rituxan is clinical regimens that employ six or eight infusions of rituximab, rather than four, a protocol that was first tested in the early clinical trials in 1995.[38] Rituxan is also being used for patients who are newly diagnosed with lymphoma, as opposed to patients who are refractory to chemotherapy.[43–47] Rituxan is now being used as a maintenance therapy, where a course of rituximab (four infusions) has been given every 6 months for a period of 2 years. This use has been tested, and clinical data suggest that the maintenance therapy is safe and efficacious.[45–47] Finally, Rituxan is being used off-label at very high doses for treatment of chronic lymphocytic leukemia (CLL) and in combination with chemotherapeutic regimens.[48,49]

The impressive (although not statistically significant because of the small sample size of the trials) results of combining rituximab and chemotherapy for low-grade follicular NHL are now being investigated in several large clinical trials.[50–52] There are several patients still in complete remission for more than 9 years after they first received chemotherapy and rituximab in 1994. More impressively, in a large, controlled trial comparing chemotherapy with chemotherapy plus rituximab for treatment of aggressive NHL (diffuse large B-cell lymphoma), the combination of rituximab and chemotherapy has been shown to improve survival in elderly patients.[53,54] Rituxan was the first drug in 20 years to extend survival in this patient population.

In addition to treating tumors, and because depleting B cells in humans with rituximab has been found to be safe, Rituxan is now in clinical trials and has shown efficacy for a number of autoimmune disorders. These include rheumatoid arthritis,[55–58] idiopathic thrombocytopenia purpura (ITP, a platelet-depletion disorder),[59,60] and lupus.[61] It is expected that depleting B cells with rituximab will be approved in a nononcology indication within several years.

32.4 Conclusions

The development and commercial success of the first monoclonal antibody used to treat tumors in the United States have had broad consequences. Companies have pushed ahead with the development of a variety of additional monoclonal antibodies. Six years after rituximab was approved, more than one third of the 38 approved recombinant proteins manufactured in mammalian cells were antibodies (some of the most notable are listed in Table 32.1). Many of these antibodies are used at large doses that were not thought to be economically practical 10 years ago. Today, the pharmaceutical industry is facing a projected shortage of manufacturing plants for recombinant proteins synthesized by mammalian cells, and there are still more than 400 monoclonal antibodies reported to be in clinical trials.

Is the discovery and development of another blockbuster monoclonal antibody like rituximab a feasible research and development model today? Many of the special events that contributed to its success are probably not replicable. As with several other successful drugs, the market turned out to be much larger than anticipated, so competition from other drugs being simultaneously developed was not so intense. Similarly, the ability to manufacture the drug in a cost-effective manner did not exist when the project was initiated, but there were other advantages unique to rituximab. Its toxicity is minimal, a

TABLE 32.1
Notable Monoclonal Antibodies Approved

Drug Name	Generic Name	Company	Antigen/ Antibody Receptor	Indication	FDA Approval
Orthoclone OKT3	Muromonab-CD3	Ortho Biotech, Bridgewater, NJ	Anti-CD3	Transplants	1986
ReoPro	Abciximab	Centocor, a division of Johnson& Johnson New Brunswick, NJ	Anti-IIb/IIIa	Angioplasty	1994
Rituxan	Rituximab	IDEC (now Biogen Idec), Cambridge, MA	Anti-CD20	Non-Hodgkin's lymphoma	1997
Zenapax	Daclizumab	Hoffmann-La Roche Nutley, NJ	Anti-IL2	Transplants	1997
Herceptin	Trastuzumab	Genentech, South San Francisco, CA	Anti-HER2	Breast cancer	1998
Simulect	Basiliximab	Novartis, Basel, Switzerland	Anti-IL2	Transplants	1998
Synagis	Palivizumab	MedImmune, Gaithersburg, MD	Anti-RSV	Infection	1998
Remicade	Infliximab	Centocor	Anti-TNF	Rheumatoid arthritis	1999
Mylotarg	Gemtuzumab ozogamicin	Ayeth Wyerst, Madison, NJ	Chemically conjugated anti-CD33	Adult myelogenous leukemia	2000
CamPath	Alemtuzumab	ILEX , Pharmaceuticals Millenium, Cambrige, MA	Anti-CD52	Chronic lymphocytic leukemia	2001
Humira	Adalimumab	Abbott Laboratories, North Chicago, IL	Anti-TNF	Rheumatoid arthritis	2002
Zevalin	Ibritumomab tiuxetan	Biogen Idec	Radiolabeled anti-CD20	Non-Hodgkin's lymphoma	2002
Bexxar	Tositumomab	Corixa, Seattle, WA (originally Coulter Pharmaceuticals) and GlaxoSmithKline, Middlesex, England	Radiolabeled anti-CD20	Non-Hodgkin's lymphoma	2003
Xolair	Omalizumab	Genentech, Novartis, and Tanox (Houston, TX),	Anti-IgE	Allergic asthma	2003
Raptiva	*Efalizumab*	Genentech, Xoma, Berkeley, CA	Anti-CD11a	Psoriasis	2003
Erbitux	Cetuximab	Bristol Myers Squibb, New York, NY/Imclone, New York, NY	Anti-EGFR	Colon cancer	2004
Avastin	Bevacizumab	Genentech	Anti-VEGF	Breast cancer	2004
Tysabri	Natalizumab	Biogen Idec	Anti-alpha 4	Multiple sclerosis	2005

Note: IL2, interluekin 2; HER2, human epidermal growth factor receptor 2; RSV, respiratory syncytial virus; TNF, tumor necrosis factor; IgE, Immunoglobulin E; EGFR, epidermal growth factor receptor; VEGF, vascular endothelial growth factor.

distinct advantage for off-label indications, especially because the drug was not anticipated to be nearly as safe, because of its depletion of normal B cells, as it appears to be. Off-label use is much larger than use in the indication that the FDA approved (treatment of refractory low-grade NHL). Development was exceptionally rapid, especially for a small company developing their first successful drug, taking only 7 years from discovery to approval. In part, this development was rapid because the nonblinded, single-arm, pivotal trial only required 166 individuals because refractory low-grade NHL was an orphan indication. For another biotech company to find a drug whose anticipated market was small enough to avoid competitors and gain important federal regulatory and financial benefits (orphan status) and yet have enough off-label potential to fuel expansive growth, it would have to find the needle in the haystack. Would that company then martial all of its resources, risk financial ruin, anticipate relevant restrictions years in advance of their adoption, and abandon other products in its pipeline, all to secure FDA approval and get the drug to market? Well, that sounds like an once-in-a-lifetime opportunity.

Acknowledgments

Many of the Rituxan team members are not acknowledged individually in the text, so I will acknowledge them now. Special thanks to John Leonard for his steady consul and tremendous effort; Alice Wei for her regulatory acumen, Chris Burman for his manufacturing expertise; Kathy Allen for her project management skills and friendship; and Lisa Franco for her process science skills. Rituxan would not have been possible without the guidance and leadership of both Bill Rastetter and Nabil Hanna. Finally, I must acknowledge that the reason I came to IDEC in 1990 was to fulfill a promise to my wife, Miyo, for us to return to California where she grew up and where we met in 1976. And lastly, a special thanks to my eldest son, Jeremy, whose writing skills I can only dream about.

References

1. Kohler, G. and Milstein, C., Continuous cultures of fused cells secreting antibody of predefined specificity, *Nature*, 256, 495–497, 1975.
2. Nadler, L.M., Stashenko, P., Hardy, R., Kaplan, W.D., Button, L.N., Kufe, D.W., Antman, K.H., and Schlossman, S.F., Serotherapy of a patient with a monoclonal antibody directed against a human lymphoma-associated antigen, *Cancer Res.*, 40, 3147–3154, 1980.
3. Morrison, S.L., Johnson, M.J., Herzenberg, L.A., and Oi, V.T., Chimeric human antibody molecules: mouse antigen-binding domains with human constant region domains, *Immunology*, 81, 6851–6855, 1984.
4. Newman, R., Alberts, J., Anderson, D., Carner, K., Heard, C., Norton, F., Raab, R., Reff, M., Shuey, S., and Hanna, N., "Primitization" of recombinant antibodies for immunotherapy of human diseases: a macaque/human chimeric antibody against human CD4, *Biotechnology*, 10, 1455–1460, 1992.
5. Weidle, U.H., Borgya, A., Mattes, R., Lenz, H., and Buckel, P., Reconstitution of functionally active antibody directed against creatine kinase from separately expressed heavy and light chains in non-lymphoid cells, *Gene*, 51, 21–29, 1987.
6. Cattaneo, A. and Neuberger, M.S., Polymeric immunoglobulin M is secreted by transfectants of non-lymphoid cells in the absence of immunoglobulin J chain, *EMBO J.*, 6, 2753–2758, 1987.

7. de Waele, P., Feys, V., van de Voorde, A., Molemans, F., and Fiers, W., Expression in non-lymphoid cells of mouse recombinant immunoglobulin directed against the tumor marker human placental alkaline phosphatase, *Eur. J. Biochem.*, 176, 287–295, 1988.
8. Vandamme, A.M., Bulens, F., Bernar, H., Nelles, L., Lijnen, R.H., and Collen, D., Construction and characterization of recombinant murine monoclonal antibody directed against human fibrin fragment-D dimer, *Eur. J. Biochem.*, 192, 767–775, 1990.
9. Wood, C.R., Dorner, A.J., Morris, G.E., Alderman, E.M., Wilson, D., O'Hara, R.M., and Kaufman, R.J., High level synthesis of immunoglobulins in Chinese hamster ovary cells, *J. Immunol.*, 145, 3011–3016, 1990.
10. Sheeley, D.M., Merrill, B.M., and Taylor, L.C.E., Characterization of monoclonal antibody glycosylation: comparison of expression systems and identification of terminal α-linked galactose, *Anal. Biochem.*, 247, 102–110, 1997.
11. Lifely, M.R., Hale, C., Boyce, S., Keen, M.J., and Phillips, J., Glycosylation and biological activity of CAMPATH-1H expressed in different cell lines and grown under different culture conditions, *Glycobiology*, 5, 813–822, 1995.
12. Reff, M.E., Adams, K., and Anderson, D.R., Comparison of recombinant chimeric immunoglobulin secreted and purified from Chinese hamster ovary (CHO) and murine SP2/0 myeloma cells, in *American Chemical Society Conference: Recovery of Biological Products VII*, American Chemical Society, San Diego, CA, September 25–30, 1994, p. 26.
13. Reff, M.E., Carner, K., Chambers, K.S., Chinn, P.C., Leonard, J.E., Raab, R., Newman, R.A., Hanna, N., and Anderson, D.R., Depletion of B cells in vivo by a chimeric mouse human monoclonal antibody to CD20, *Blood*, 83, 435–445, 1994.
14. Liu, A.Y., Robinson, R.R., Murray, E.D., Jr., Ledbetter, J.A., Hellstrom, I., and Hellstrom, K.E., Production of a mouse-human chimeric monoclonal antibody to CD20 with potent Fc-dependent biologic activity, *J. Immunol.*, 139, 3521–3526, 1987.
15. Urlaub, G., Mitchell, P.J., Kas, E., Chasin, L.A., Funanage, V.L., Myoda, T.T., and Hamlin, J., Effect of gamma rays at the dihydrofolate reductase locus: deletions and inversions, *Somat. Cell Mol. Genet.*, 12, 555–556, 1986.
16. Davies, J. and Reff, M., Chromosome localization and gene-copy-number quantification of three random integrations in Chinese-hamster ovary cells and their amplified cell lines using fluorescence in situ hybridization, *Biotechnol. Appl. Biochem.*, 33, 99–105, 2001.
17. Ringold, G., Dieckmann, B., and Lee, F., Co-expression and amplification of dihydrofolate reductase cDNA and the *Escherichia coli* XGPRT gene in Chinese hamster ovary cells, *J. Mol. Appl. Genet.*, 1, 165–175, 1981.
18. Davies, J., Jiang, L., Pan, L.-Z., LaBarre, M., Anderson, D., and Reff, M., Expression of GnTIII in a recombinant anti-CD20 CHO production cell line: expression of antibodies with altered glycoforms leads to an increase in ADCC through higher affinity for FcγRIII, *Biotechnol. Bioeng.*, 74, 288–294, 2001.
19. Miller, R.A., Maloney, D.G., Warnke, R., and Levy, R., Treatment of B-Cell lymphoma with monoclonal idiotype antibody, *N. Engl. J. Med.*, 306, 517–522, 1982.
20. Stashenko, P., Nadler, L.M., Hardy, R., and Schlossman, S.F., Characterization of a human B lymphocyte-specific antigen, *J. Immunol.*, 125, 1678–1685, 1980.
21. Manshouri, T., Do, K.A., Wang, X., Giles, F.J., O'Brien, S.M., Saffer, H., Thomas, D., Jilani, I., Kantarjian, H.M., Keating, M.J., and Albitar, M., Circulating CD20 is detectable in the plasma of patients with chronic lymphocytic leukemia and is of prognostic significance, *Blood*, 101, 2507–2513, 2003.
22. Jilani, I., O'Brien, S., Manshuri, T., Thomas, D.A., Thomazy, V.A., Imam, M., Naeem, S., Verstovsek, S., Kantarjian, H., Giles, F., Keating, M., and Albitar, M., Transient down-modulation of CD20 by rituximab in patients with chronic lymphocytic leukemia, *Blood*, 102, 3514–3520, 2003.
23. O'Keefe, T.L., Williams, G.T., Davies, S.L., and Neuberger, M.S., Mice carrying a CD20 gene disruption, *Immunogenetics*, 48, 125–132, 1998.
24. Shan, D., Ledbetter, J.A., and Press, O.W., Signaling events involved in anti-CD20-induced apoptosis of malignant human B cells, *Cancer Immunol. Immunother.*, 48, 673–683, 2000.

25. Demidem, A., Lam, T., Alas, S., Hariharan, K., Hanna, N., and Bonavida, B., Chimeric anti-CD20 (IDEC-C2B8) monoclonal antibody sensitizes a B cell lymphoma cell line to cell killing by cytotoxic drugs, *Cancer Biother. Radiopharm*., 12, 177–186, 1997.
26. Polyak, M. and Deans, J., Alanine-170 and proline-172 are critical determinants for extracellular CD20 epitopes: heterogeneity in the fine specificity of CD20 monoclonal antibodies is defined by additional requirements imposed by both amino acid sequence and quaternary structure, *Blood*, 99, 3256–3262, 2002.
27. Chan, H.T., Hughes, D., French, R.R., Tutt, A.L., Walshe, C.A., Teeling, J.L., Glennie, M.J., and Cragg, M.S., CD20-induced lymphoma cell death is independent of both caspases and its redistribution into triton X-100 insoluble membrane rafts, *Cancer Res*., 63, 5480–5489, 2003.
28. Cragg, M.S., Morgan, S.M., Chan, H.T., Morgan, B.P., Filatov, A.V., Johnson, P.W., French, R.R., and Glennie, M.J., Complement-mediated lysis by anti-CD20 MAb correlates with segregation into lipid rafts, *Blood*, 101, 1045–1052, 2003.
29. Cartron, G., Dacheux, L., Salles, G., Solal-Celigny, P., Bardos, P., Colombat, P., and Watier, H., Therapeutic activity of humanized anti-CD20 monoclonal antibody and polymorphism in IgG Fc receptor FcγRIIIa gene, *Blood*, 99, 754–758, 2002.
30. Clynes, R.A., Towers, T.L., Presta, L.G., and Ravetch, J.V., Inhibitory Fc receptors modulate in vivo cytotoxicity against tumor targets, *Nat. Med*., 6, 443–446, 2000.
31. Harjunpaa, A., Junnikkala, S., and Meri, S., Rituximab (anti-CD20) therapy of B-cell lymphomas: direct complement killing is superior to cellular effector mechanisms, *Scand. J. Immunol*., 51, 635–641, 2000.
32. Treon, S., Mitsiades, C., Mitsiades, N., Young, G., Doss, D., Schlossman, R., and Anderson, K., Tumor cell expression of CD59 is associated with resistance to CD20 serotherapy in patients with B-Cell malignancies, *J. Immunother*., 24, 263–271, 2001.
33. Weng, W.K. and Levy, R., Expression of complement inhibitors CD46, CD55, and CD59 on tumor cells does not predict clinical outcome after rituximab treatment in follicular non-Hodgkin lymphoma, *Blood*, 98, 1352–1357, 2001.
34. Byrd, J., Kitada, S., Flinn, I., Aron, J., Pearson, M., Lucas, D., and Reed, J.C., The mechanism of tumor cell clearance by rituximab leukemia: evidence of caspase activation and apoptosis induction, *Blood*, 99, 1038–1043, 2002.
35. Anderson, D.R., Grillo-Lopez, A., Varns, C., Chambers, K.S., and Nanna, N., Targeting cytotoxic immunotherapy, *Biochem. Soc. Trans*., 25, 705–708, 1997.
36. Grillo-Lopez, A., White, C., Dallaire, B., Varns, C., Shen, C., Wei, A., Leonard, J., McClure, A., Weaver, R., Cairelli, S., and Rosenberg, J., Rituximab: an insider's historical perspective, *Semin. Oncol*., 27, 9–16, 2000.
37. Czuczman, M., Grillo-Lopez, A.J., White, C.A., Saleh, M., Gordon, L., LoBuglio, A.F., Jonas, C., Dallaire, B., and Varns, C., IDEC-C2B8/CHOP chemoimmunotherapy in patients with low grade lymphoma: clinical and BCL-2 (PCR) results, *J. Mol. Med*., 75, B231, 1997.
38. Grillo-Lopez, A.J., White, C.A., Varns, C., Shen, D., Wei, A., McClure, A., and Dallaire, B.K., Overview of the clinical development of Rituximab: first monoclonal antibody approved for the treatment of Lymphoma, *Semin. Oncol.*, 26(suppl. 14), 66–73, 1999.
39. Maloney, D.G., Grillo-Lopez, A.J., White, C.A., Bodkin, D., Schilder, R.J., Neidhart, J.A., Janakiraman, N., Foon, K.A., Liles, T.-M., Dallaire, B.K., Wey, K., Royston, I., Davis, T., and Levy, R., IDEC-C2B8 (Rituximab) anti-CD20 monoclonal antibody therapy in patients with relapsed low-grade non-Hodgkin's lymphoma, *Blood*, 90, 2188–2195, 1997.
40. Grillo-Lopez, A., White, C., Dallaire, B., Varns, C., Shen, C., Wei, A., Leonard, J., McClure, A., Weaver, R., Cairelli, S., and Rosenberg, J., Rituximab: the first monoclonal antibody approved for the treatment of lymphoma, *Curr. Pharm. Biotechnol*., 1, 1–9, 2000.
41. Manz, R.A., Thiel, A., and Radbruch, A., Lifetime of plasma cells in the bone marrow, *Nature*, 388, 133–134, 1997.
42. Foran, J., Norton, A.J., Micallef, I., Taussig, D., Amess, J., Rohatiner, A., and Lister, T.A., Loss of CD20 expression following treatment with Rituximab (chimeric monoclonal anti-CD20): a retrospective cohort analysis, *Br. J. Haematol*., 114, 881–883, 2001.

43. Hainsworth, J.D., Burris III, H.A., Morrissey, L.H., Litchy, S., Scullin, D.C., Jr., Bearden III, J.D., Richards, P., and Greco, F.A., Rituximab monoclonal antibody as initial systemic therapy for patients with low-grade non-Hodgkin's lymphoma, *Blood*, 95, 3052–3056, 2000.
44. Hainsworth, J., Rituximab as first-line systemic therapy for patients with low-grade lymphoma, *Semin. Oncol.*, 27, 25–29, 2000.
45. Hainsworth, J., Rituximab as first-line and maintenance therapy for patients with indolent non-Hodgkin's lymphoma: interim follow-up of a multicenter phase II trial, *Semin. Oncol.*, 29(Suppl. 2), 25–29, 2002.
46. Hainsworth, J.D., First-line and maintenance treatment with rituximab for patients with indolent non-Hodgkin's lymphoma, *Semin. Oncol.*, 30(Suppl. 2), 9–15, 2003.
47. Hainsworth, J.D., Litchy, S., Barton, J.H., Houston, G.A., Hermann, R.C., Bradof, J.E., and Greco, F.A., Single-agent rituximab as first-line and maintenance treatment for patients with chronic lymphocytic leukemia or small lymphocytic lymphoma: a phase II trial of the Minnie Pearl Cancer Research Network, *J. Clin. Oncol.*, 21, 1746–2751, 2003.
48. Keating, M. and O'Brien, S., High-dose rituximab therapy in chronic lymphocytic leukemia, *Semin. Oncol.*, 27, 86–90, 2000.
49. Lin, T.S., Lucas, M.S., and Byrd, J.C., Rituximab in B-cell chronic lymphocytic leukemia, *Semin. Oncol.*, 30, 483–492, 2003.
50. Czuczman, M., Fallon, A., Mohr, A., Stewart, C., Bernstein, Z., McCarthy, P., Skipper, M., Brown, K., Miller, K., Wentling, D., Klippenstein, D., Loud, P., Rock, M., Benyunes, M., Grillo-Lopez, A., and Bernstein, S., Rituximab in combination with CHOP or fludarabine in low-grade lymphoma, *Semin. Oncol.*, 29(Suppl. 2), 36–40, 2002.
51. Hiddemann, W., Dreyling, M., and Unterhalt, M., Rituximab plus chemotherapy in follicular and mantle cell lymphomas, *Semin. Oncol.*, 30(Suppl. 2), 16–20, 2003.
52. Van Oers, M.H., Hagenbeek, A., Van Glabbeke, M., and Teodorovic, I., Chimeric anti-CD20 monoclonal antibody (Mabthera) in remission induction and maintenance treatment of relapsed follicular non-Hodgkin's lymphoma: a phase III randomized clinical trial — Intergroup Collaborative Study, *Ann. Hematol.*, 81, 553–557, 2002.
53. Coiffier, B., Rituximab in combination with CHOP improves survival in elderly patients with aggressive non-Hodgkin's lymphoma, *Semin. Oncol.*, 29, 18–22, 2002.
54. Coiffier, B., Immunochemotherapy: the new standard in aggressive non-Hodgkin's lymphoma in the elderly, *Semin. Oncol.*, 30(Suppl. 2), 21–27, 2003.
55. Protheroe, A., Edwards, J.C.W., Simmons, A., Maclennan, K., and Selby, P., Remission of inflammatory arthropathy in association with anti-CD20 therapy for non-Hodgkin's lymphoma, *Rheumatology (Oxford)*, 38, 1150–1152, 1999.
56. Edwards, J.C.W. and Cambridge, G., Sustained improvement in rheumatoid arthritis following a protocol designed to deplete B lymphocytes, *Rheumatology (Oxford)*, 40, 205–211, 2001.
57. Leandro, M.J., Edwards, J.C., and Cambridge, G., Clinical outcome in 22 patients with rheumatoid arthritis treated with B lymphocyte depletion, *Ann. Rheum. Dis.*, 61, 883–888, 2002.
58. Cambridge, G., Leandro, M.J., Edwards, J.C., Ehrenstein, M.R., Salden, M., Bodman-Smith, M., and Webster, A.D., Serologic changes following B lymphocyte depletion therapy for rheumatoid arthritis, *Arthritis Rheum.*, 48, 2146–2154, 2003.
59. Zaja, F., Iacona, I., Masolini, P., Russo, D., Sperotto, A., Prosdocimo, S., Patriarca, F., de Vita, S., Regazzi, M., Baccarani, M., and Fanin, R., B-cell depletion with rituximab as treatment for immune hemolytic anemia and chronic thrombocytopenia, *Haematologica*, 87, 189–195, 2002.
60. Zaja, F., Vianelli, N., Sperotto, A., De Vita, S., Iacona, I., Zaccaria, A., Masolini, P., Tomadini, V., Tani, M., Molinari, A.L., Baccarani, M., and Fanin, R., B-cell compartment as the selective target for the treatment of immune thrombocytopenias, *Haematologica*, 88, 538–546, 2003.
61. Eisenberg, R., SLE — Rituximab in lupus, *Arthritis Res. Ther.*, 5, 157–159, 2003.

33

Funding the Birth of a Drug: Lessons from the Sell Side

Howard E. Greene

CONTENTS

33.1 Introduction

The pharmaceutical industry is hugely capital-intensive. It now costs in the range of $1.5 billion to commercialize the average new drug.* Most of these investments will never pay off. Historically, less than a third of new drugs have generated sufficient revenue to cover their development costs.† And yet, astonishingly, drug discovery is a hotbed of entrepreneurial activity and new company startups.

This chapter is intended to provide insight into the process of funding drug discovery and development in startup companies that are years from generating positive cash flows, if ever. Its target audience is people who are interested in joining the biotech arena, but who have no experience with the unique sources of funding that sustain it. My aim is to describe lessons I learned from the "Sell Side" of Wall Street transactions, i.e., entrepreneurs who by selling their stock are trying to build company value, as opposed to investors who by buying the stock are trying to build portfolio value, the "Buy Side." The goal is to cover topics that cannot be learned at business schools, from INC Magazine, or in venture capital brochures.

33.2 Background

Since the lessons to be learned here are from one man's life experiences, they should be viewed in the context of my career and the biases it created. My career started out in a

*http://www.bain.com/bainweb/publications/printer_ready.asp?id=14243
†http://www.phrma.org/publications/publications/brochure/questions/whycostmuch.cfm

conventional manner: physics degree from Amherst College, MBA from Harvard University, 7 years with McKinsey & Company, an executive position with Baxter International. Then, as Bill Graham, the legendary Chairman of Baxter, years later put it so well, "The best thing we ever did for Ted was to fire him." Bill's next sentence reflected what my flameout from corporate America taught me to do well: turning to Baxter's CEO, he concluded, "If Baxter ever runs short of money, call Ted."

At the time of my departure from Baxter in 1978, I did not know it, but I was starting what would later be termed a "biotechnology" company. Nor, did I realize that I would spend most of the next two decades raising money to shovel into the black hole of biotech discovery and development. The ensuing roller coaster left me with strong opinions, including certain biases which, while not always shared by the financial establishment, are usually echoed by my colleagues who also worked the Sell Side. In addition to having a hand in nine startup companies, I co-founded a successful venture capital firm and served on the board of a biotech mutual fund; these two Buy Side experiences reinforced my biases expressed here.

The biotech business combines three extremely satisfying activities: pioneering breakthrough science, meeting important social needs, and achieving huge financial returns. But, the technical risks are not for the faint of heart. Biotech is for the 21st century what wildcat oil drilling was for the 20th century; in both industries, a company can do everything exactly right, but end up with a dry hole. Being smart can definitely shift the odds in favor of success, but it is impossible to know if there is oil down there without drilling a very deep, expensive hole. In our lifetimes pharmaceutical discovery and development will continue to be a wildcat venture, right up until the fat lady (FDA) sings. Before anyone commits their fortune and honor to this business, they must truly understand this reality.

33.3 Defining an Investment Concept: The Story Is Everything

The first dictionary definition of the word "story" is: "An account or recital of an event or a series of events, either true or fictitious."[‡] In the investment community, biotech business plans are referred to as "stories," presumably because they combine both truth (they are sincerely made proposals of future reality) and fiction (they have not happened yet, and probably would not).[§] So, get used to this word; it is not a pejorative reference to your business plan.

The financial heart of a biotech story is a positive risk-adjusted net present value (rNPV). If the projected payback could happen very quickly, and if investors regarded a business plan as a sure thing (risk-free), then the "value" of the story would simply be the payback minus the investment. However, biotech paybacks do not happen quickly and are certainly not risk free. Hence, investors discount the payback value to reflect: (1) the time value of their money and (2) the risk of failure. If discounted payback values are thought to exceed the required investments, stories get funded; if this rNPV is perceived as negative, the flow of money stops.

If you do not thoroughly understand the concept of rNPV, then you simply cannot understand why anyone would dedicate $1.5 billion to a 15 year project that is likely to

[‡] http://www.dictionary.com

[§] Among investors, biotech companies have also come to be known as "deals." In my opinion, this terminology reflects the unfortunate evolution of the VC industry away from comprising entrepreneurs who had started and built their own companies and toward pure financial managers for whom the flesh and blood enterprises amount to not much more than IRR calculations.

prove disappointing. If rNPV is not firmly in your financial toolbox, stop right here, buy a good book on financial principles (i.e., *Principles of Corporate Finance* by Brealey and Myers), study the rNPV chapter, then resume reading this. (For a short course, see Putting a price on biotechnology in *Nature Magazine* by Stewart, Allison, and Johnson[¶] — what to do with?

33.3.1 Positive Biotechnology Investment Strategies

It is important to understand the quantitative theory of rNPV, because it provides insight into the elements that contribute to a positive investment decision. In the case of biotech, two components of a technology strategy usually overshadow all other variables, the sales potential and the risk of failure.

33.3.1.1 Sales Potential

The valuation of successful drugs is so high that profit projections are essentially irrelevant. At this writing (early 2004), Amgen was trading at a stock price above ten times trailing 12-month (TTM) revenues. By comparison, General Motors was trading below eight times TTM earnings, which resulted in a stock price of only about one eighth of TTM revenues. In other words, per dollar of sales, drugs are worth 80 times cars and trucks.

33.3.1.2 Risk of Failure

"Of every 5000 medicines tested, on average, only 5 are tested in clinical trials and only 1 of those is approved for patient use."[‖] Very, very few biotech concepts yield commercial returns. Consequently, investors are hypersensitive to events which appear to shift the odds of success one way or the other. What drives biotech stock prices up or down is not changes in timelines, but rather changes in perceptions of the risk of failure.

The two key elements of any biotech story are the importance of the medical need and the rationale for why this idea is likely to work when previous efforts have failed. While the time value of money should theoretically be important, given the long timelines for drug development, in practice estimates of successful drug valuations and technical risks usually overshadow the risk-free time cost of money.

33.3.2 Biotechnology Stories

To simplify their financial models, over the years professional investors have developed at least four categories of biotech stories.

33.3.2.1 Product Stories

These plans envision composition-of-matter patents on drug molecules for which development timelines and market potentials can be estimated. Product stories are the clearest, most rational plans because profitable sales can be projected and rNPVs calculated. On the other hand, investors usually perceive higher risks because the rNPVs tend to be dominated by single product candidates for which historical failure rates are high. Examples of product stories include Amgen's development of Epogen and Amylin's development of Symlin.

[¶] http://www.nature.com/cgi-taf/DynaPage.taf?file=/nbt/journal/v19/n9/full/nbt0901-813.html&filetype=pdf
[‖] http://www.phrma.org/publications/publications/brochure/questions/whycostmuch.cfm

33.3.2.2 Biology Stories

These plans envision pioneering new fields of science in order to reap the financial rewards due pioneers. Companies may or may not start out with method-of-use patents on biological mechanisms that can be targeted or utilized for disease treatments. It is usually not possible to estimate rNPVs early in the story, but the concept of "broad technology platforms" is beloved by professional investors who view such platforms as hedging their bets with multiple disease targets. While the view of ultimate commercial payoffs is blurred, so are the harsh realities of drug development risk. Also, such biology stories usually have enthusiastic backing from academics who have already boarded the relevant NIH funding train. A brilliant example of this strategy was the exploitation of numerous genomic discovery companies, who drafted in the academic slipstream behind the Human Genome Project.

33.3.2.3 Toolbox Stories

These plans envision method-of-use patents on ways to discover new drugs. They follow a pattern of first making money selling discovery services to established pharmaceutical companies, to be followed by proprietary drugs developed for the companies' own accounts using toolbox revenues. Examples of such stories include Medarex and Abgenix, who patented technologies for creating antigen-specific human monoclonal antibodies; Vertex and Agouron, who focused on rational drug design technologies; and Aurora, who developed an innovative high-throughput screening system.

33.3.2.4 Development Stories

These plans envision assembling a portfolio of in-licensed drug candidates for which collective development risk is deemed to be relatively low. The idea is to in-license molecules that either are too specialized (small market potential) for large pharma or are not visible to their radar because they come from places like Eastern Europe. These stories are favorites of professional investors, because rNPVs can be calculated with relatively low development risk on the basis of demonstrated clinical utility and/or pipeline diversity. Examples include Dura Pharmaceuticals, who marketed prescription products that treat infectious and respiratory diseases, and Gilead who marketed antiviral nucleotides discovered in the Czech Republic.

33.4 Universal Truths of the Buy Side

As investor sentiment waxes and wanes — along with successes and failures among companies — the relative popularity of these biotech categories changes. There are, however, a few universal truths which underlie all business plans and which enhance investor enthusiasm and occasionally cause Buy Side stampedes.

33.4.1 New Stories Trump Reality

Investors love new ideas, because they do not have any of the warts and scars that inevitably collect on tested pharmaceutical concepts. People who invest in biotech are looking for medical breakthroughs that can go where no drug molecule has gone before. Perhaps this enthusiasm derives from the natural hope that, finally, a cure for cancer or diabetes or atherosclerosis will be found. It certainly taps the same optimism that the

charitable organizations depend upon. Meanwhile, the inverse of this phenomenon also applies: old stories that suffered setbacks, even if they provided insight into a route through the development minefield, are viewed as road kill by most investors.

33.4.2 Big Ideas Attract Big Investors

The term "Big Idea" was coined for me by Don Valentine, who is among the venture capital legends of Silicon Valley. What he means is that, especially at the earliest stage, Big Ideas generally take no more effort on the part of investors than do small ideas, and are, therefore, strongly preferred. Big Ideas also enhance venture capital exits. During one biotech initial public offering (IPO), a senior manager of a large, well-known mutual fund told me that he did not care about reviewing the clinical data, because the product's market potential was so huge that he was buying every share he could get his hands on. This was a classic example of how really Big Ideas can sometimes preempt scientific reality.

33.4.3 Risk Is a Dirty Word

It has always amazed me that "portfolio managers" insist on each company taking steps to minimize their individual corporate business risks. As any investment book explains, the basic rationale for owning a portfolio is to spread the risk associated with each holding. Thus, it is the job of the portfolio manager, not each of his holdings, to spread the risk. The reality is that, in the business of startup companies, the only way small companies can succeed is to focus, focus, focus. As Sir Richard Sykes, now retired CEO of Glaxo, said to me when I proposed diversifying Amylin's pipeline, "Pharmaceuticals are among the most difficult businesses on this planet, and focus is the key to success." At that time, as I recall, the bulk of his entire $1+ billion Research and Development (R&D) budget was focused on seven key projects. Unfortunately, even at the venture capital stage, most professional investors are very concerned about taking write-offs, given the politics of staying employed in their business, i.e., their "fiduciary responsibility." So, the most popular biotech stories often incorporate unrealistic illusions, i.e., "broadly enabling technologies," to camouflage the fact that if their main project blows up, the company is probably finished.

33.4.4 Biotech Depends Upon Other People's Money

No single investor has the stomach for the risk and time required to take a drug idea all the way to market. Thus, professional investors expect that other people's money (OPM) will either (1) absorb significant amounts of financial risk or (2) provide an exit strategy at higher valuations. Biotech business plans that attract private stage investors emphasize OPM to mitigate negative burn rates: NIH grants, big pharma license fees and development support, and IPOs. Among public investors, OPM reduces the risk of running out of money, but more importantly it is the sincerest form of endorsement when the cash comes from sources who are presumed to be technologically sophisticated.

33.5 Endorsements Trump Science

Most professional investors are clueless about the science underlying biotech companies (and, they would be quick to say so). As a consequence, who endorses a biotech plan usually

carries more weight than what supports it in terms of data and logic. These endorsements flow from three sorts of "experts" at the top of their respective food chains.

33.5.1 Distinguished Academic Scientists

I have always thought it curious that investors would think that people who chose not to pursue a career in business can provide reliable advice about the value of a business plan. Moreover, in most fields of science the major breakthroughs, which change the course of human events, are initially dismissed by the academic elite who have built their careers on the old ways of thinking. I call this the "Copernicus Effect."** Nevertheless, biotech entrepreneurs looking for funding from professional investors are well advised to solicit friendly reviews from Nobel Laureates or potential nominees. I have racked my brain, but I cannot remember a situation where a famous, senior-statesman academic actually contributed meaningfully to the success of an early-stage biotech company, other than via endorsements.*† The real help seems to always come from the younger academics who have no stake in the status quo, and who are willing to take intellectual risks.

33.5.2 Recognized Business Leaders

Investors are usually eager to bet on successful track records in business ventures. Senior executives from large pharmaceutical companies can attract money, even though they may not be able to articulate the science, and probably never led a small research-oriented team while raising money. Even more attractive are executives who have already created value in startup companies that made newsworthy profits for their investors. For entrepreneurs without these qualifications, such business leaders can improve the quality of a board of directors while being a magnet for money. But, remember, a 30-year career rising though the ranks of large companies does not prepare one well for the bone-chilling nakedness of a small, startup company.

33.5.3 Extremely Successful Investors

Money attracts money. By far, the best endorsements for attracting new investors are the financial commitments of professional investors who have made noteworthy fortunes in the biotech arena. Usually, these are venture capitalists who have learned to judge entrepreneurs and their stories, or at least have been very lucky. Herein lies one of the main reasons for inviting high priced VC money into your deal: it is a very good investment for attracting more OPM in subsequent funding rounds.

To summarize, a good biotech story has the following elements: it is a new, really Big Idea that mitigates the risk with "broadly enabling technology," it falls in a business category that is currently hot among investors, it proposes to depend primarily on OPM through deal making, and it has got an all star cast (a Nobel Laureate in biochemistry, a

** Readers not familiar with the story of how Nicholas Copernicus disrupted 1000 years of astronomical wisdom according to Ptolemy should study this history. Copernicus got it right with his heliocentric model of the solar system, but any "due diligence" by senior academics of his time would have concluded he was dead wrong. All academic establishments have a stake in maintaining the validity and relevance of the work that got them to the top of their professions.

*† Having said this, I must pay special tribute to Richard Lerner and Sidney Brenner, whose counsel and support greased the tracks for my launching scientific endeavors that flew in the face of conventional wisdom.

retired CEO of Merck or a founder of Amgen, and the senior biotech partner of a premier venture capital firm).

If your story has all these elements, there is one additional factor that will affect your ability to raise money and that is timing. Investors' appetites for biotech stories are like the wind: if it is blowing hard enough, any kite will fly; as the wind dies down, only the very best kites will fly; and occasionally, during a dead calm, no kites will fly. And, as with the wind, mere mortals cannot control these cycles, they can only wait for the wind to improve.

Having a great story is essential to financing a biotech company. Once that is in place, it is critical to understand the various sources of biotech funding.

33.6 Understanding Sources of Capital: Making a Deal with the Devil

Unless you have inherited $1.5 billion, the only way you can successfully bring a biotech drug to market is to use OPM. This section focuses on the various sources of OPM, with some thoughts about their relative desirability.

A question I am consistently asked about funding a biotech company is, "What kind of investor can afford to accept the extremely high technical risk of drug development while waiting 10 to 15 years for a product launch?"

The answer is, no single investor can get the job done. Rather, funding a biotech company requires a chain letter strategy, whereby subsequent groups of investors achieve their target returns by passing the stock "baton" to the next level of investors as technical risks decline.

33.6.1 Four Primary Investor Categories

33.6.1.1 Seed Round Investors

Seed round investors focus on converting raw scientific data into a viable business plan. This category of funding can come from a wide variety of sources, including individual entrepreneurs, angel investors, academic licensing departments, corporations spinning out technology, and venture capital firms. Ideally, at this stage the core management team is recruited. Seed round investors achieve their return when companies go public, hopefully 5 to 7 years after startup.

33.6.1.2 Private Investors

Private investors focus on the high risk, low burn-rate portion of the business plan, usually aimed at carrying the company through the Phase 2 evidence of a relevant therapeutic effect in man. These investors, usually venture capital firms, tend to fixate on their "exit strategy" for achieving liquidity, either by selling the company or by taking it public, ideally 3 to 5 years after investing.

33.6.1.3 Corporate Investors

Corporate investors are interested in acquiring rights to marketable drugs, so they "invest" in biotech companies by paying licensing fees, covering research expenses, awarding milestone payments, or buying stock. Most companies today are reluctant to make major commitments until after Phase 2 data are available, so their timelines to product launch are in the 5 to 10 year range.

33.6.1.4 Public Investors

Public investors purchase stock in IPOs or follow-on rounds after the technical risk has declined and the stock is liquid. Some of these investors expect to make their return when important commercial milestones are achieved, i.e., positive Phase 3 results, FDA approval, or financial breakeven, usually within a few years. Other investors are "momentum players" who hope to buy the stock on an upswing and get out before investor enthusiasm wanes, often within several months. Still other speculators enjoy playing the high volatility of biotech stocks, hoping to make returns within several days.

If the chain letter is successful, the end result is a profitable company valued on the basis of conventional investment criteria, i.e., price/earnings ratio. During the 15 to 20 years after founding that is typically required to achieve profitability, a biotech company will have had a constantly changing shareholder list. With the exception of a few visionary founders, these investors expect to make their money over relatively short time frames.

33.6.2 Investor Subgroups

Within the early investor categories lie various subgroups of investors, which have their advantages and disadvantages for biotech companies.

33.6.2.1 Family and Friends

This is usually cheap but stupid money. Valuations are rarely considerations, so stock can often be sold at higher prices than would be acceptable to professional investors. Unless Bill Gates is your cousin, funding amounts are usually very small. And, when the going gets tough, this group of investors can offer little business help. More importantly, they may become a burden when they start to worry about their investments. I have never seen this source of OPM be very helpful over the long run.

33.6.2.2 Angel Investors

These are investors who have made their money in the biotech industry and wish to recycle some in new ventures. Angel investors understand the risks and expect only a small fraction of their early-stage investments to achieve a good return. Often they can provide advice or introductions useful to the venture. In my experience, angel investors are, by far, the best group of investors because they expect to lose most of their money. But, they are rarely deep pocketed enough to step in when the program requires serious funding.

33.6.2.3 Venture Capital

This is the most expensive, but smartest money for Big Idea biotech startups. Recruiting the right VCs as shareholders brings expertise in building companies and usually very deep pockets. In this regard, certain firms are known as investment leaders in the biotech field, and their individual partners can serve as useful board members. The trick is finding VCs who want to build companies at least as much as they want to make money, because company-builders are, unfortunately, a very small percentage of the breed.*‡ Over the years, I have had the privilege of working with a number of company-builder VCs

*‡ While it is not a perfect indicator, it sometimes helps to keep track of how often VCs refer to their portfolio companies as "deals." If they never describe their investments in scientific or human terms, they are pure money managers with no personal commitment to the success or failure of the ventures.

(i.e., Tom Perkins and Jim Blair), who stuck with management through thick and thin because they were personally committed to the ventures' successes. It is unfortunate that "until death do us part" is not a universal preamble to VC financing agreements.

33.6.2.4 Pharmaceutical Companies

Inevitably, a biotech venture must turn to big pharma for funding. Such deals are essential for validating the commercial potential of the technology, and the bigger companies always contribute important resources that are beyond the reach of small startups. Unfortunately, however, big pharma's commitment to individual projects is very low, since most of them fail and strategic priorities change over time. Projects can be canceled for a variety of reasons not related to the technical or commercial merits of a program. Mergers present a huge risk to biotech deals, because the first postmerger activity is to arrange the merged R&D programs into a ranking according to strategic priorities; more often than not, this ranking will delete high-risk biotech projects. In one painful experience, our program was dropped because big pharma needed a few extra cents per share, and canceling us was the expeditious way to achieve this bottom line effect. Unfortunately, Wall Street is completely naive about these corporate realities, so, even though valuable technology rights are returned, investors often stampede for the exit when a deal is terminated.

33.6.2.5 Institutional Investors

Any biotech company wanting to accomplish more than just cashing royalty checks must ultimately turn to public investors for funds. Going public provides the VC investors a return on their investment, and only the public markets can provide funding in the 100 million dollar ranges. Unfortunately, only a small number of institutional investors display any long-term commitment to Big Ideas. Most of these Buy Side investors do not understand the technology or business, so they put their money into stocks that they think promise upward price "momentum." Because of their fiduciary responsibility to their own investors, institutions tend to be very risk averse; they would rather pay a high price for a safe stock with modest upside than a low price for a risky stock with great potential.

33.6.2.6 Retail Investors

Last on the list of OPM are the individuals who invest their own money on Wall Street. Rarely do they play a significant role in company financing, since their individual investments are too small to permit efficient fund raising. The only biotech companies where retail investors represent a majority of shareholders are those that have been abandoned by institutional investors, i.e., the road kill.

33.6.3 Key Fundraising Concepts

These groups and subgroups represent the most important sources of OPM for biotech companies. When addressing these sources, it is important to understand two key concepts that apply universally to fundraising.

33.6.3.1 Kleiner's First Law: "If You Can Get It, Take It"

Gene Kleiner was one of the truly wise men of Silicon Valley, having co-founded Fairchild Semiconductor, as well as Kleiner & Perkins, an incredibly successful venture capital firm. What Kleiner's First Law teaches is that biotech companies should raise capital whenever they can, assuming the deals do not have fatal flaws (as will be discussed below). I once

violated Kleiner's First Law, when one of our public offerings was about to be priced below the threshold I considered reasonable. Based on concern about dilution, I called off the deal. We subsequently ran short of cash and had to accept far worse dilution than would have been caused by the deal I rejected.

This taught me in spades that biotech entrepreneurs who are fixated on minimizing dilution of their shares risk company death by financial starvation. When a modest amount of dilution is the principal downside of a proposed financing, "If you can get it, take it."

A corollary of Kleiner's First Law is, "Raise money when you do not need it." If you wait until your need for cash is desperate, you can be sure the financial terms will be unattractive. The less you need the money, the more leverage you will have with investors. In fact, a general principle to keep in mind is that investors love to put their money where it is not needed for survival, but rather is needed to expand the upside potential of the business plan.

33.6.3.2 Good Lead Investors Are Key to Good Deals

After years of struggling to raise money for biotech companies, I have concluded that every private deal must have a lead investor. The "lead" is the individual (and his firm) whose commitment persuades other investors to participate. Some combination of two qualities designate the lead investor: (1) recognized expertise and success in the field which makes the deal credible; and (2) sufficient financial commitment to the deal to get it done without most of the other investors.

The majority of investors are herd animals who follow the smart money. Every private deal I have ever done ultimately came together when a lead investor emerged who could drive the other investors into the offering corral. Put another way, most investors will board the funding train only when they know it is about to leave the station without them.

33.7 Rules of the Game: Raising Money Is an Art Form

Understanding who has the money and how they operate is the necessary foundation for funding a biotech operation. But this knowledge is useless if you do not understand the rules of the game.

In my experience, the most successful biotech entrepreneurs are those who are driven by the medical and commercial potential of their science, but who never forget that their earliest "product" is paper. Biotech ideas need cash, and cash comes from doing deals.

Most professional investors will confess to having only a vague understanding of biotechnology and medicine, but they all consider themselves experts in deal structuring. It is what they do for a living. Thus, the biotech CEO who dreams of changing the practice of medicine must first understand the art of deal making.

33.7.1 The Art of Valuation

The central issue in any stock offering is price. In traditional, profitable companies price is usually measured as a price/earnings ratio. Since biotech companies rarely are profitable, price is evaluated using the capitalized value of the outstanding stock, i.e., price per share times the total number of shares, compared to other companies at similar stages of development with comparable upside potential. This "market cap" number (either private or public) is what sophisticated biotech investors look to in measuring whether an offering price is fair. Two measures used are postmoney and premoney values.

33.7.1.1 Postmoney Value

Postmoney value is the market cap after the funding is complete and is what new investors focus on. Postmoney value is measured against other deals and the likely step-up in value prior to the next funding event. As the biotech market heats up or cools down, comparable values rise and fall. During the venture capital stage, investors try to estimate when the company can do an IPO (their exit strategy) and at what premoney value, and based on this estimate, coupled with their target rate of return, they discount back to what would be a reasonable private postmoney value.

33.7.1.2 Premoney Value

Premoney value is the market cap before the funding is complete and is what existing (old) investors focus on. If the company has met its milestones with no downturn in public valuations then the premoney value is usually higher than the postmoney value of the previous funding rounds. This is a happy event for existing investors and employees who hold stock or options. If progress has been disappointing, or biotech valuations in general are down, premoney value may be less than the previous round, which is very disappointing for stockholders. The emotional impact of a "down-round" can be very hard on early-stage companies.

Bottom line, valuation in biotech companies is analogous to bidding on eBay: if there are more buyers for a deal than stock available, premoney value rises; if there is a shortage of buyers, premoney value falls. The key to doing a strong deal is to orchestrate a shortage of stock and an excess of buyers.

33.7.2 Less Is More

Investors are always most anxious to buy what they cannot have. If investors view an attractively valued stock to be in short supply, they will bid it up in price and in amount of cash they are willing to invest. Hence, it is vitally important to create the notion that there is not enough stock to go around. To do this, biotech management must orchestrate the perception of scarcity, and they must clearly (and privately) differentiate between the amount of cash they are aiming to raise (the budget) and the amount of cash investors think they are willing to raise (the script). In most cases these are two different numbers, the first for strictly internal consumption and the second for publication to new investors. The goal is to create a deal that is "over-subscribed," i.e., where demand exceeds supply. In private deals, over-subscription can increase demand at the next funding round; in public deals, over-subscription ensures there is postdeal demand for the stock that will absorb the downward price pressure from investors who "flip" their stock for a quick profit. The trick is to know when deal size has reached the point where opening the spigot any further will begin to jeopardize investor enthusiasm, and this is where having the most skillful investors and advisors on board really pays dividends.*§

33.7.3 Avoiding Live Hand Grenades

Traditionally, in startup ventures shareholders who supply cash occupy a "preferred" position over shareholders who supply intellectual property (IP) and blood-sweat-and-tears,

*§ This may appear to violate Kleiner's First Law. In fact, over-subscribed deals always raise more money than was originally offered, and, by expanding the deals, managements are following KFL.

i.e., investors own preferred stock while founders and employees own common stock. The preferred stock converts to common at an IPO or acquisition. The aim of this arrangement is to return to investors at least a portion of the remaining financial value when a company fails and is wound up. Unfortunately over the past decade, VCs (and some institutional investors in public stocks) have gone beyond this basic liquidation protection and have devised ways to shift more of the future valuation risk to common shareholders. The two most egregious practices include full ratchets and participating preferred.

33.7.3.1 Full Ratchets

Full ratchets are antidilution provisions that are aimed at preventing a decline in the value of investors' stock when a down-round occurs, i.e., when stock price declines. Usually, at company startup the conversion ratio between preferred and common stock is set at 1:1. Traditionally, this ratio has been subject to revision on a weighted-average basis when down-rounds occur. For example, if investors own 50% of the company, then the preferred conversion ratio is adjusted to reflect a price that is half way between the original preferred price and the new round preferred price. This adjustment, called a "weighted-average ratchet," shifts some of the value lost in a down-round from the preferred to the common shareholders (i.e., punishes founders and employees). During the 1990s, some VCs began requiring "full ratchets," which means the preferred was repriced to exactly match the new, lower price. Not only is a full ratchet severely punishing to shareholders who put up the intellectual and sweat capital, but it can cause a "death spiral" if the new, premoney valuation goes low enough. As the old preferred conversion ratio rises, so also does the resulting number of postconversion common shares, and this accelerates the fall in stock price on a premoney basis. If the premoney value falls far enough, the common shareholders are wiped out. If a funding source proposes a full ratchet, run, do not walk, to the door.

33.7.3.2 Participating Preferred

Participating preferred is another mechanism aimed at shifting risk from the VCs to company founders and employees. The idea here is that, if the company is sold before going public, the VCs get back their entire investment before the common shareholders share pro-rata in any remaining value. Under particularly punishing versions of participating preferred, the VCs are entitled to receiving multiples of their investment, say 2X, before the common shareholders share in any remaining value. The plain vanilla version is not, in my opinion, particularly fair to the folks who took the most risk, but it does demotivate them from liquidating the company for a share of its residual financial value. On the other hand, preference terms with payback multiples greater than one reflect the kind of ethics that caused Jesus to throw the money changers out of the temple.

The thing to remember is read the fine print in the term sheet. Do not be dazzled by a high premoney valuation if one of these hand grenades is built into the deal. If you permit them in, you can be sure that every subsequent round of investors will insist on equal terms, and the risk to common shareholders will escalate.

33.7.4 Risks to Common Shareholders

33.7.4.1 Pay to Play

Always keep in mind that those friendly, smiling VCs who will provide you money and serve on your board are competitors. When they depart your boardroom, they will gladly knife each other in the back. Their continued survival and lifestyle depends upon raising future rounds of money from investors who can choose from alternative VC funds on the

basis of their past profitability, i.e., internal rate of return. As a result their most vicious behavior occurs when the going gets tough for a company and one of their fellow investors decides to walk away. "Over their dead body," will they allow the walk-away to achieve a decent return without risking more capital. To spike this traitor, the remaining VCs will engineer a "crunch round" at a very, very low price, which wipes out the non-players — along with the founders and employees. Since employees are usually essential to building a company, the remaining VCs will issue new stock or options to restore their motivation. However, they will arbitrarily decide who should come through the process with any upside value, and the results of this process are not pretty. Be sure you focus on investors who have a reputation for long-term commitment to their companies, as well as deep pockets, and cross your fingers that they all will continue to pay to play.

33.7.4.2 Price Validation

Unfortunately, VCs are required to keep track from quarter to quarter of the value of their portfolio. If all the stocks were publicly traded, this would be simple. The problem is that most VC stock holdings are privately held, so their prices must be arbitrarily estimated, with the most straightforward assumption being the price last paid by investors. However, this assumption could lead to false valuations if a fixed group of VCs simply continued investing in the same company at ever higher prices, which would mathematically, if not really, increase the value of their earlier holdings. Hence, the industry has a practice of bringing new investors into each financing round to "validate" the new price on the basis of arms length negotiation. The downside of this process, of course, is that at every funding cycle management must drop everything to seek out, educate, and negotiate with new investors. In recent years, many VCs have tried to streamline this process by putting very large sums of money up front with the intention of carrying the company through key clinical milestones, thereby obviating the need for interim price validation. The downside of this approach is that companies who start out with too much money tend to be wasteful. Damned if you do, damned if you do not.

33.7.4.3 Price Indifference

Every biotech CEO needs to understand the mathematics of dilution. Never worry about the deal price without calculating what the true dilutive impact is between their price and your price. Usually the impact is vanishingly small, and other deal points, e.g., full ratchets, are much more important. Do not let ego get in the way of price negotiations. Also, run the numbers to show yourself that any investor who buys the same percentage of the deal as he already owns in the company is completely indifferent to the final price. If the price is low, he buys a higher share of the company with his new money; if the price is high, then he retains a higher share of the company with his old shares. This is the "sweet spot," and the best VCs stay in it by taking their pro-rata share with each succeeding round.

33.7.4.4 The Exit Strategy

As discussed earlier, it is a rare investor who can hold on to a biotech stock for the decades it takes to achieve the full upside business potential. Thus, at each stage of company development, each group of investors is looking toward selling their stock to new investors whose risk/reward criteria match the next stage. The most obvious transition occurs when a company does an IPO, because VC firms are structured to distribute their stock holdings to their investors (general and limited partners) when that stock becomes liquid, i.e., salable on NASDAQ. An alternative liquidity event is sale of the private company to a larger company for cash or public stock. For this reason, biotech CEOs need to get used to their

investors and board members fixating on "exiting" the company, and it is important to be sensitive to this need. As insulting as it seems, it comes with their money.

33.7.4.5 Be Ready for Windows

Because biotech valuations are usually based on perceptions of future value rather than present earnings and dividends, stock prices are subject to wild fluctuations as investors alternate between optimism (sometimes euphoria) and pessimism (sometimes despair). In a rising market, stock can be sold; in a falling market, stock cannot be sold. The times when investors have an appetite for biotech stocks are termed "financing windows." They are impossible to predict with certainty, and by the time it is obvious the window is open, it is usually starting to close. Thus, it is important to be ready to jump when the window opens. During the private company phase, investors' appetites tend to parallel the public markets', since they perceive better exit opportunities. At the moment of an IPO, it is important to have all the ducks lined up: the legal, financial, managerial aspects plus a banker ready to launch. After the IPO, follow-on offerings should be done when the opportunities present themselves, not according to some long-term financing plan. Never forget Kleiner's First Law: "If you can get it, take it."

The path from concept to drug meanders through a financial forest, with wild beasts roaming just beyond sight. Hopefully, the concepts covered here will assist the biotech entrepreneur in staying on the path and continuing to move forward.

33.8 Living with the Consequences: What You Need Once You Start

Here are some general principles that are useful to keep in mind while trying to build a biotech company.

33.8.1 No Intellectual Property, No Money

Given the time and money required to bring a drug to market, IP is absolutely essential to providing a fair return on investment. Biotech companies can get started without owning patents, but very soon after their founding they simply must invent or in-license patent-pending ideas. In the pharmaceutical industry patent protection is absolutely essential for protecting the profit margins. Personal skills or trade secrets will not attract the magnitude of investment required to get to market. Large companies will not do major development deals unless they can access important IP, primarily because they feel competent to meet any technical challenge within their spheres of expertise.

33.8.2 People Are More Important than Science

Big pharma companies are inevitably market-driven, and if science does not represent an obstacle for them, they can get the job done. Thus, to succeed commercially, biotech companies must be science-driven. Nevertheless, in my experience it takes outstanding people to successfully pioneer a new field of science. At the beginning, endorsements from recognized superstars in relevant fields are useful for attracting funding. Even more importantly, in areas of biology where no road maps exist, it takes a special breed of persons to hack their way through the forest, both in terms of intelligence and perseverance. Bottom

line, it takes people to make ideas work. For this reason, always aim to recruit people who are smarter than yourself.

33.8.3 Budgeting under Uncertainty

No business can operate without budgets, but there are special budgeting problems for biotech companies. First, cash inflows are almost impossible to forecast, given the uncertainty of timing and size. Second, expense plans can change abruptly on the basis of new data that forces a change in direction. As a result, 12-month "budgets" are really "forecasts" subject to considerable change. In this environment the standard corporate budgeting process does not work very well. What is needed are rolling 12-month budgets, which are updated quarter-by-quarter.

33.8.4 From Hip-Shooter to Brain Surgeon

During the early stages of a pure, science-driven biotech venture, the leadership qualities that drive success include understanding the technology, dreaming-up commercial strategies and deal structures, articulating the story, and promoting the upside to employees, partners, and investors. Organizational structure should be minimal, with close lines of communication and a real team spirit. However, as a company progresses into manufacturing, clinical trials, and marketing, more traditional methods of pharmaceutical management become important. In my experience, rarely is the individual who successfully launched the company into orbit the correct person to run it once its orbital velocity has been achieved. Somewhere about the start of Phase 3 trials, it is usually time for the founding CEO to pass the baton to a real pharmaceutical executive.

33.8.5 Staying the Course

Creating a first-in-class drug is among the most difficult challenges in modern capitalism. The CEO of a biotech firm needs to be comfortable in complex science, intellectual property, medical practices, government regulations, and Wall Street finance for his company to succeed. Since almost no academic or job training can prepare someone for this spectrum of expertise, the only way is to plunge in and start swimming. And, it is a long, hard swim to reach the goal, so it takes a mighty big dreamer to make it worthwhile. But that is what the biotech industry is made of, so get out there and go for it.

33.9 Conclusion

Half a millennium ago Christopher Columbus set out on a voyage that changed the world forever. The academic establishment of his time believed it was a foolish idea to sail west, because India was so far away his crew would starve before they arrived. And, even if he got there by sailing down wind, how could he get home against the wind? In fact, from a scientific standpoint they were absolutely correct. But, from a commercial standpoint, they were dead wrong about the value of Columbus' voyage.

The most successful biotech voyages have similar characteristics. They are launched in pursuit of drug ideas that most experts in big pharma consider too remote and risky. Moreover, some of the most successful drug ideas have directly contradicted the academic wisdom of

their time. These pioneering ideas have been championed by individuals who refused to listen to the counsel of supposed experts who were certain that a particular idea would not work.

One of my favorite anecdotes was told to me by Jan Leshly, who at the time was President of SmithKline Beecham. He had brought together a distinguished group of academic scientists to review a project of questionable potential. At the end of the day, he called for a vote, and the academics unanimously declared that the project should be terminated. After dismissing the committee, Jan turned to his management team and said, "So, it is clear what we must do. Full speed ahead with the project!"

In other words, often the most dramatic breakthroughs have come from sailing upwind against the current wisdom. Unfortunately, sources of money rarely have sufficient expertise on their own to make judgments about technical risks, so their wind blows in the same direction as academic opinion. Thus, the real drug pioneers face a daunting task, and they are pretty much on their own, just as Columbus was.

But, in the end, if a controversial drug idea can be brought to fruition, the result is among the most satisfying outcomes in modern society. Great drug ideas can do enormous good for the human condition, and individuals who make meaningful contributions to pharmaceutical research can finish their careers with an enormous sense of satisfaction. And, given the financial rewards of a successful, proprietary drug, this satisfaction extends beyond the spiritual to the material awards of a prosperous venture.

On the other hand, if a drug idea does not work, blessed are those who gave their all to see it through. In this regard, I think the following quote from Teddy Roosevelt nicely sums up the thrill of drug discovery:

> It is not the critic who counts; not the man who points out how the strong man stumbled, or where the doer of deeds could have done them better. The credit belongs to the man who is actually in the arena, whose face is marred by dust and sweat and blood; who strives valiantly; who errs and comes short again and again; who knows the great enthusiasms, the great devotions; who spends himself in a worthy cause; who, at the best, knows in the end the triumph of high achievement, and who, at the worst, if he fails, at least fails while daring greatly, so that his place shall never be with those timid souls who know neither victory or defeat. (*Citizenship in a Republic: The Man in the Arena.* http://www.geocities.com/rainforest/3745/tr.html)

34

Innovations for the Drug Development Pathway: What Is Needed Now

Janet Woodcock

CONTENTS

34.1 Rising Expectations

There are rising public expectations about the prospects for new therapies based on new biomedical discoveries. There has been tremendous investment over the last 30 years by the National Institutes of Health (NIH) and the private sector in basic biomedical science. The sequencing of the human genome has been accomplished and new technologies have arisen from that progress. There are also advances in medical imaging and the advent of nanotechnology and its applications in medicine, as well as innovations in tissue engineering. And on the discovery side, new chemical synthetic techniques and automated screening have dramatically increased the number of early candidates available for evaluation.

34.1.1 Increased Spending

During the past decade, there has been an increased push to invest in biomedicine. The NIH funding doubled starting in about 1998 and pharmaceutical R&D has been increasing since then. There have been major investments in biotechnology. Figure 34.1 shows indexed growth. On the top is pharmaceutical R&D spending; in the triangles below are NIH budget increases. The increases have been in parallel.

34.1.2 Reduced Review Time

The FDA review times used to be considered as a major bottleneck for drug development, but this problem has diminished tremendously (Figure 34.2). The "drug lag" of the 1980s and early 1990s has disappeared. In general, FDA review time for priority (the most novel) drugs is very prompt and hovers around 6 months. For standard drugs, it varies a little from year to year because the numbers of new candidates are low, but total time for approval takes about a year and a half.

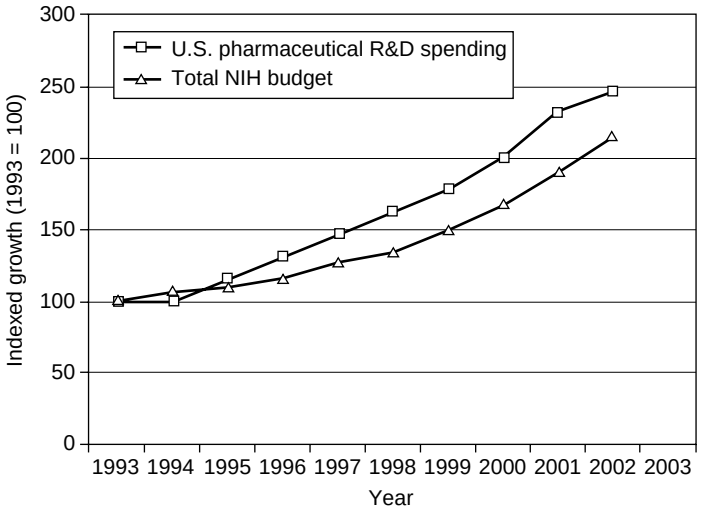


FIGURE 34.1
NIH funding and R&D spending.

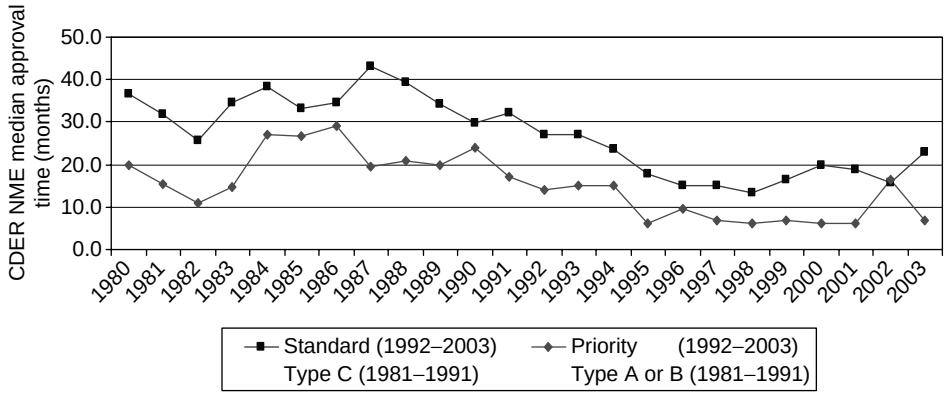


FIGURE 34.2
Review time. Prior to 1992, therapeutic gain was classified as type A, B, or C. Starting in 1992, priority and standard designation was used to represent therapeutic potential for new drug approvals.

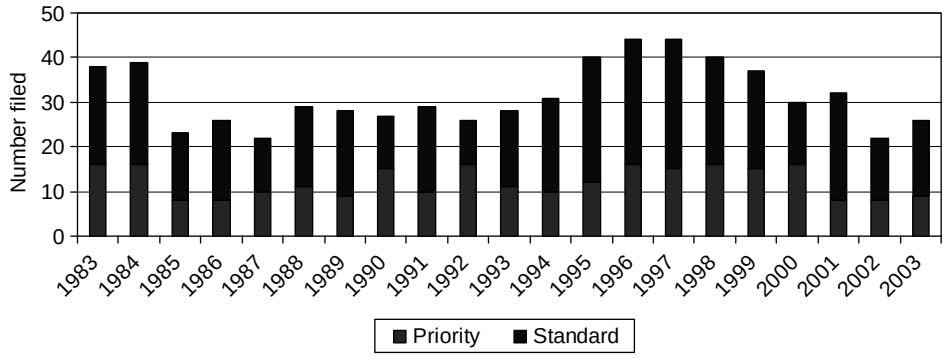


FIGURE 34.3
New product submissions. For NMEs submitted prior to 1992, type A and type B applications were counted as priority review and type C applications were counted as standard review.

34.2 Indicators of Drug Development Problems

With all these positive indicators, the expectation has been that a large number of new products will become available. That was the belief in the mid-1990s. But Figure 34.3 shows what has happened over time: New product submissions have basically remained flat for the last 20 years. Figure 34.3 depicts new molecular entities submitted to the FDA since 1983, and there is no trend toward increasing submissions. In fact, concern has been raised lately about the downturn since the late 1990s. Whether or not this is a real trend or simply a fluctuation is not known; the point is that there has not been a huge increase in the submission of new products.

Clearly, there is a plateau or a dip in the new drug pipeline. The cause is multifactorial and has been the subject of debate in the literature. Many people think that genomics and other new sciences are not at their full potential for contributing yet; it often takes 10 to 15 years for fundamental new discoveries to be effectively used by innovators and be translated into new technologies. In addition, the plethora of new genomically based targets that have been uncovered have caused a temporary decrease in candidate drugs because of confusion about the large number of targets and the inability to distinguish which are viable.

Other people believe that mergers and other business arrangements in the pharmaceutical industry have decreased the number of candidates. Still others opine that the easy targets have been taken or that drug development is now targeting more difficult diseases — chronic neurodegenerative disease, for example. Finally, some hold that the rapidly escalating costs and complexity decrease pharmaceutical companies' willingness or ability to bring candidates forward into the clinic. In fact, there is probably truth in all of these explanations, and there are probably additional pertinent explanations.

34.2.1 Increasing Development Costs

Whatever the explanation for the above problems, it is true that development costs per new product are now higher. Although the number in Figure 34.4 can be disputed, depending on which assumptions are incorporated, the cost of clinical drug development has escalated rapidly in the past 5 years. For common diseases and treatments, many companies find that the cost of a single efficacy trial may approach $100 million. The details can be debated, but most people agree that the investment that is (not simply the direct costs) required to bring a successful candidate into the market is somewhere between $800 million and the estimate in Figure 34.4 of US$1.7 billion. Obviously, there will be products on the low end of that investment, for orphan drugs or diseases with smaller populations, but there are going to be products that go through a torturous development process and are at the high end of this investment figure. So, the cost of bringing a new drug to market is somewhere in the hundreds of millions of dollars.

34.2.2 Increasing Failure Rate

In addition to rising costs, for some reason, new candidates now have a lower chance of success in the clinic. This finding has been validated using many methods. Although varying figures are quoted in the literature, the pharmaceutical industry generally agrees with the following assertion: Despite all the advances in science, the success rate of product development — our ability to identify successful candidates — has not improved, in fact, it has deteriorated. It is estimated that new compounds entering phase 1 have about an 8% chance of

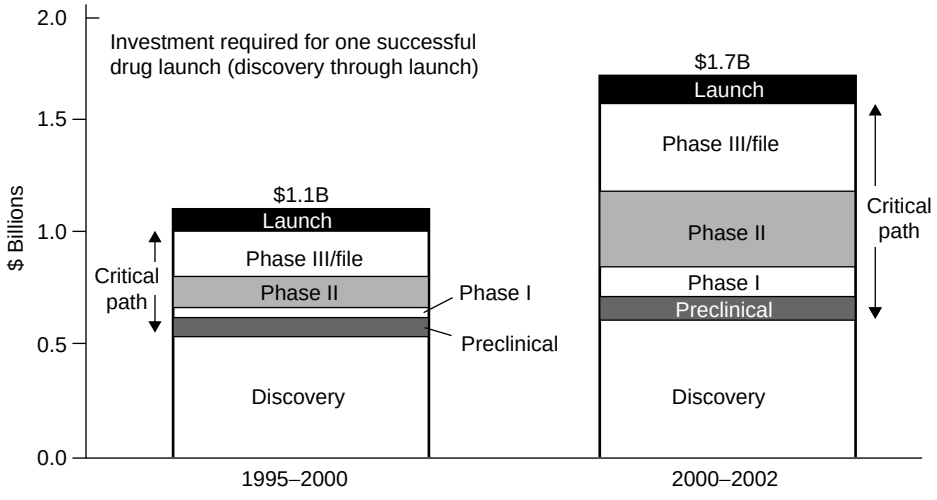


FIGURE 34.4
Increasing development costs. (From the Business & Medicine Report, 2003).

success vs. about a 14% chance of success 15 years ago. That is not a huge deterioration, perhaps not a statistically significant deterioration; however, it is going in the wrong direction.

Of compounds that enter into the very expensive phase 3 trials, the failure rate is now estimated to be about 50% vs. an estimated 20%, 10 years ago. Again, there are a litany of explanations for this: for example, some developers are more reckless because they want to get into phase 3 more quickly because of the huge costs, or that people are attempting to treat much more challenging disorders. Whatever the explanation, which again is probably multifactorial, these figures describe the predictability of the process right now. Yet, almost no company enters a compound into the clinic without very high hopes for it. Unfortunately, the luster of any compound decreases the more it is studied. This lowering of the drug development success rate, despite all the investment in science, is a great concern, not only because of the absolute magnitude of the costs, but also because this price tag decreases the system's ability to study many candidates.

If this trend continues, therefore, it will have a negative effect on the public health. I think we are already seeing some of this. Because of the costs and barriers to entry into the market, companies must focus their effort primarily on revenue blockbusters to get the return on investment required to keep the process going. Clinically important therapies that do not promise the largest revenues may take a backseat to products that have a higher chance of a positive return. We have seen, for example, antibiotics programs closing down, and we are concerned about the development of vaccines, medical countermeasures, and blood products; but primarily, we are concerned about the development of targeted therapy and drug targets in a smaller disease population. How attractive will these programs be if the investment costs to bring a product to market are so high? We, as a society, need more innovation in areas where the public health need is most urgent. The increasing costs of development create a very high bar to competition and to the entry of smaller firms into the very innovative areas, and this also will not benefit the public.

34.3 Enhancing Systemic Understanding

Many people say that the answer to these problems is more investment in basic science, and that is true; no one disagrees with that assertion. Eventually, predictability of the development process will improve as basic science evolves. Biological knowledge will have to progress from single-pathway understanding to complex, interactive-system understanding: achieving that level of understanding will take many decades. Not only normal physiology, but also pathophysiology will need to be understood at the systems level. Such knowledge will take us toward a true mechanistic understanding of therapeutic intervention, which will enable us to design therapies based on first principles.

Currently, we can build an airplane on a computer or design a bridge using all the data available and engineers do not expect those designs to fail 92% of the time after they are built. The enterprise of drug development is infinitely more risky. The engineers and the basic scientists can design airplanes on a computer because they have knowledge linking to the first principles and through the applied sciences: enough relevant information that they can build a new airplane that is almost always going to fly.

34.3.1 The Societal Challenge

While we in drug development are waiting for this level of mechanistic understanding to evolve, a societal challenge continues: There is urgent need for medical product innovation

in many areas now. For instance, the epidemic of inactivity and obesity in this country is going to accelerate illness and the burden of illness, rapidly, through all age groups of the population, not just the aging of the "baby boomers," but the premature aging of younger populations, based on sedentary and extremely overweight lifestyles. That will create a tremendous challenge to the healthcare system. And this system is currently overburdened and in crisis.

34.3.2 Targeted Therapy

We also need targeted therapy. The failure of candidate and marketed drugs that have to be withdrawn shows that we still do not have enough information about them. When drugs are approved, we basically know that they work, and we know they are safe for a short period of time. That is about all the information we are able to glean from the very expensive drug development process. Often, we understand a drug's pharmacology, but we do not understand its dose parameters very well. We certainly do not understand, to any great extent, who should take this drug and who should not take it, either as that relates to effectiveness or to enhanced safety.

We need to get a lot better at this. This is one of the most concerning aspects of the current drug development process: It does not yield that much information at the end of the day, from a clinical perspective, about how to use the drug. Yet the process is nevertheless expensive and cumbersome, and the cost, during the past decade, has increased tremendously. This is what is fueling the desire to import unapproved drugs from Canada and other countries; people are beside themselves about the cost of their medicines. The reimbursers, who pay for drugs, need a higher level of certainty about a drug's performance, both its safety parameters and its effectiveness, which is just not available from current methods of evaluation. There is growing skepticism that our investment in basic biomedical research will deliver on its promise. Society did not invest in basic research for the pursuit of pure knowledge; it invested in basic research to improve human health.

34.4 Food and Drug Administration Diagnosis: The Critical Path to New Medicines

The FDA examined these drug development challenges, and on March 16, 2004, issued the report: *Innovation/Stagnation: Challenge and Opportunity on the Critical Path to New Medical Products*, and created an initiative. The "Critical Path Initiative" constitutes FDA's diagnosis of certain drug development problems and a prescription for what to do about them. The Critical Path Initiative is a serious attempt to bring attention and focus to the need for targeted scientific efforts to modernize the techniques and methods used to evaluate the safety, efficacy, and quality of medical products as they move from product selection and design to mass production.

34.4.1 The Applied Science of Development

Although I have mentioned other barriers and problems in drug development and some of the reasons for them, FDA thinks there might be one underlying problem that is addressable now, that does not require waiting decades for basic biomedical science to become truly predictive. We believe that societal investment in the research and development

efforts to improve the drug development process has been lacking relative to the huge private and public investment in basic research. Society has invested a large amount of money and focus on genomics and related science. No one is saying that was inappropriate; we would not have learned about basic biology or achieved the current level of understanding of drugs, if we had not done that — but investment in the middle piece, the applied science, that in building bridges or designing airplanes would be considered engineering that has been lacking. There has not been an appreciation of need for the applied science to enhance the science of development at the same time that the number of candidates and pathways and targets has exploded.

34.4.2 The Applied Science Gap

Academia is not funded, overtly, to engage in research into the applied science of development. That has also never been conceptualized as an FDA role because FDA is viewed, by government and the public, as a regulatory agency. The fact that FDA has to use science to compose its regulations and standards, and that the science for that has to come from somewhere, has not been recognized. What about the private sector? On the one hand, people thought, "Well, NIH does all this work," and on the other, people said, "Well, the private sector is supposed to take the basic science and develop it into products."

In fact, each of these sectors has a different role. The NIH contributes basic biological and disease knowledge, and the private sector discovers and develops medical innovations. The Critical Path Initiative, however, is not about discovery nor is it about developing specific drugs. These are separate activities. Everyone recognizes that improvement in new technology, perhaps even some new paradigms, is needed in these areas, but they are not FDA's area of expertise. The NIH is getting more involved in the translational realm with the "Roadmap Initiative." They are trying to establish, for example, tissue banks, repositories, and other tools that would aid in discovery activities. Critical Path, in contrast, is concerned with the applied science work: after a candidate, is discovered, what is needed to move it to the market? What other kinds of science need to be applied?

The FDA thinks this applied science gap could be addressable now with the right attention and focus. We can use new scientific knowledge and other existing technologies to improve the development process. The research community can develop robust, applied research programs into the critical path — scientific work that will lead to generalized knowledge that everyone could use: for example, acceptable new trial designs or endpoints. The academic basis for this type of applied science needs to be strengthened. Innovation, at a high level, does not happen without an academic, scientific component. The FDA's involvement in the Critical Path research and standards development should be intensified.

34.4.3 Defining the Critical Path

What does the FDA mean by the critical path? For drug development, the FDA is referring to the range of activities that start at the point where some lead candidates have been discovered and identified and proceed all the way to the point where those products get approved — that is what we are calling the Critical Path. We feel that there are sciences (Figure 34.5) underlying these activities. The literature on "translational research," usually refers to taking specific products and moving them from the bench into phase 1. Critical Path science and Critical Path research refer to the actual infrastructure: the tools used to evaluate products as they move through preclinical and clinical development. And evolving those tools is different from developing a specific product.

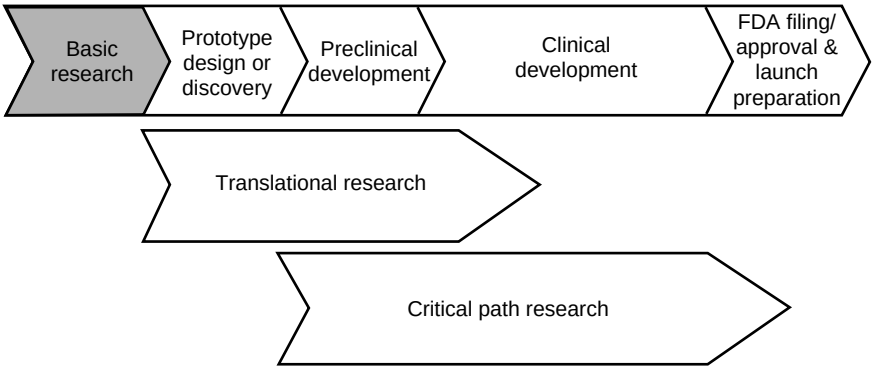


FIGURE 34.5
FDA's critical path costs.

Applied science to address three key dimensions:

- Assessment of safety – how to predict if a potential product will be harmful?
- Proof of efficacy – how to determine if a potential product will have medical benefit?
- Industrialization – how to manufacture a product at commercial scale with consistently high quality?

FIGURE 34.6
What is on the "critical path"?

34.4.4 Explaining the Process

There are three major scientific tasks required to move forward along the critical path (Figure 34.6):

1. From the beginning, a sponsor has to be able to predict that the candidate product will be relatively safe in people. It has to evaluate the candidate, both preclinically and clinically, to ensure that it has an acceptable safety profile.
2. The sponsor has to ensure that the candidate is going to benefit people as nobody is going to pay for it and the FDA will not approve it unless it is effective.
3. And finally, there is the issue of "industrialization." The critical path report encompasses all medical products: devices, biologicals, and drugs. All of these products have industrialization challenges. For example, formulating drugs that can be delivered to the site of action is an industrialization component. For a device, it might be manufacturing it in a way that is reproducible. The underlying issue of industrialization is the applied science that enables manufacturing of a product at commercial scale with consistently high quality. This is not an easy component either; many products founder because of problems in this area.

There are whole disciplines, whole industries, devoted to each of the boxes in Figure 34.6. For instance, there is a whole industry and scientific field devoted to animal toxicology.

Patient groups or other interested parties who are not familiar with the drug development process need to realize how fantastically complex this whole story is.

34.5 Collaborating to Improve the Critical Path

The FDA will be announcing shortly what, specifically, it is going to do in the upcoming year or years, for initiating the Critical Path. We take this very seriously, and we plan to move ahead. We want to set up collaborative efforts among government, academia, industry, and patient groups. The effort will encompass the applied sciences underlying development. We need to improve our investment in applied science areas and develop concrete work products, such as biomarkers, that will help everyone and provide generalized knowledge. We are not talking about developing products. Generally, government should not get into developing products. We want to build support for the basic academic science in the relevant disciplines.

Over the past several years, FDA has been working on the "Pharmaceutical Product Quality Initiative." This initiative has been directed toward improving the manufacturing of drugs and modernizing the way FDA regulates to allow for innovation and change in the manufacturing sector. This effort has required getting recognition that pharmaceutical manufacturing is one of the industrial bases. It has never been conceptualized or supported that way, but pharmaceutical manufacturing is a tremendous industry in the United States. In fact, worldwide, the pharmaceutical companies spend significantly more on manufacturing their products than they do in R&D. Nobody ever believes that statement, but the facts demonstrate that it is true. Those are real numbers from the pharmaceutical companies: They spend more on manufacturing their products than they do on R&D.

34.5.1 Generalizing and Sharing Knowledge

The cost of drug manufacturing, everyone agrees, could be significantly decreased. Manufacturing can be made much more efficient by applying modern techniques. There are analogies, therefore, with the Critical Path process. We want to build opportunities to share existing knowledge and databases. We have a tremendous amount of experience in drug development — good and bad — and one of the big problems is that those experiences are not shared. That experience, unshared, cannot be used to build generalizeable knowledge that can take us all to the next level. There are ways by which we can share knowledge and improve our general science.

34.5.2 Developing Enabling Standards

At FDA, we need to develop standards that enable drug development. The FDA received a large number of comments to our Critical Path paper: 120 different organizations submitted thoughtful comments. There were some areas of agreement. Everyone agrees that much more work needs to be done on improving clinical trials, everything from their conduct (in other words, their actual mechanics), to their infrastructure, to how they are handled, and to the science of design and analysis. There was tremendous agreement about developing better biomarkers and surrogate endpoints. In the preclinical work-up, there was a lot of discussion about enhancing predictability. There was also a separate discussion from the manufacturing and product-design side regarding product characterization and manufacturing standards.

34.5.3 Incorporating Proof-of-Mechanism Studies

In drug development, there is tremendous interest in improving early understanding of the characteristics of compounds in humans. The FDA is actually moving ahead with this now; we are going to issue guidance that will enable early proof-of-mechanisms studies to allow for screening in humans including microdose or imaging studies. Such studies would occur before initiation of the traditional drug development pathway. As humans are the most predictive species for human response, there is great interest in getting products into people, in an exploratory way, before embarking on drug development. That might mean looking at a range of candidates or a small set of candidates in people and picking the ones with the best characteristics in humans. It might mean establishing a proof-of-mechanism before investing heavily in clinical drug development. To this end, we are working on guidance on the toxicology and product characterization needed to initiate these studies. We are also working on guidance for laboratory production of clinical material. We think these efforts will enable researchers to get into the clinic more rapidly. We will issue these guidances in drafts, of course, and seek comments.

34.5.4 Improving Clinical Trial Design

Another thing that needs work is the design of trials. There has been quite a bit of talk about the use of Bayesian designs, but there has been little implementation in drug development. We need to move on and actually see some of these designs worked through to see how they perform. Trial design can also involve computer design and modeling, for example, or the enrichment designs that FDA has been promulgating. These need to be worked out at theoretical level and then tried somewhere to see if they perform as predicted.

If we are going to move toward targeted therapy, we are going to have to start using diagnostics to discriminate among various patients. Obviously, we use diagnostics now, but we need newer diagnostics and new trial designs that incorporate them. Currently, there is a reluctance to get into this area because it is innovative and might slow down the process. Yet individualization is probably the future of therapeutics.

34.5.5 Conducting Less-Expensive Trials

In the conduct of clinical trials and their underlying infrastructure, Robert M. Califf, who is at Duke and who presented at the FDA Science Board during its discussion of the Critical Path Initiative, stated that the cost of many clinical trials could be decreased by 50%, simply by streamlining and removing unnecessary steps. Later, other experienced parties stated that the 50% figure was too low, and the cost could be reduced by more than that. Whatever the actual figure is, there is a great opportunity to streamline conduct and data collection in clinical trials: part of this reform would involve using more statistically based approaches to data cleanup and handling parameters, which is an expensive part of clinical trials. Much of the current process is driven by views about FDA requirements in this area that are probably not wholly correct.

34.5.6 Standardizing the Data

We need additional progress on data standardization, including case-report forms. This effort has moved forward in the past several years through the work of the Clinical Data

Interchange Standards Consortium (CDISC) and other efforts. The push to standardize trial data must continue. Also, we are working with the FDA's bioresearch monitoring program to make sure the messages that program sends out are congruent with any new progress we make in these areas.

34.5.7 Agreeing on Clinical Trial Analyses

The FDA has generated a list of clinical trial analysis issues that need to be addressed:

- How to manage multiple endpoints statistically
- How to handle missing data
- How to incorporate enrichment designs
- How to study combination regimens
- How to design and analyze noninferiority trials

We have been discussing these for a long time. Getting a consensus guidance developed on each would constitute tremendous progress.

34.5.8 Finding and Constructing Disease-Specific Endpoints

We all realize that we need better disease-specific endpoints in very many diseases. There are tremendous opportunities to use imaging and certain biomarkers, if only they could be relied upon. Patient-centered outcome measures are also extremely important in most chronic and symptomatic diseases; yet, they are not used as extensively as they need to be. Agreement on acceptable composite endpoints is needed in many diseases.

We need to develop a better pathway for validating biomarkers and surrogate endpoints. There is great confusion about how a new biomarker can be used in a regulatory manner or how a biomarker can actually become a surrogate endpoint for approval. The biomarker process typically works like this: A quantitative measurement is discovered in someone's laboratory; the findings on this biomarker are published in the scientific literature. Then other people study it and publish their findings in the scientific literature. Finally someone starts using it, in a small clinical series, which they publish. Other people might try animal models, and publish that. The biomarker might even be applied in clinical practice, with physicians using it to monitor the course of disease, and those results are published. Then for, perhaps, a decade, that is the *status quo*, and people say, "Why isn't FDA using this biomarker as a surrogate endpoint in trials?"

The biomarker is not used because no synthetic analysis has been done. The data need to be pooled, synthesized, and analyzed. We have to understand what the data are telling us about that biomarker and what the remaining gaps in understanding are. Studies have to be identified that will fill those gaps, and then somebody has to do that work, whatever it is; and for a surrogate endpoint, of course, that work involves correlation with clinical outcomes.

The problem is: nobody is in charge of doing this. Is it academia? Is it the FDA? People come to FDA and say, "Why aren't you using this biomarker?" And the people at FDA say, "We don't have the data; nobody's analyzed the data." Private companies may try to synthesize the data and present it to the agency, but usually they do not have access to all the datasets. Those datasets reside in different companies and different hands around the country.

The purpose of a surrogate endpoint is use for regulatory approval in lieu of a clinical outcome. Surrogate endpoints have the very same set of issues as biomarkers but at a higher level. Once surrogate endpoint data are analyzed, the gap in understanding often

relates to how well that marker correlates with clinical outcomes in a therapeutic situation. To determine that, the marker has to be put into trials of successful treatments to show whether the movement of the biomarker conforms to the successful treatment of the patient, measured by clinical success. This is sometimes done, but the data are seldom synthesized. An additional step for a surrogate endpoint would be evaluation by an FDA Advisory Committee and the development of a new guidance that would advise what trial designs and what measures would need to be incorporated into such a surrogate. Safety surrogates are the same; it has to be shown to someone's comfort, to FDA's and the community's comfort, that these surrogates are actually predictive of a safety outcome.

There are many biomarkers that need this type of further development. Right now, for example, we need to improve the radiographic criteria for tumor progression. Such development has usually been done on an *ad hoc* basis in most diseases. We need to look at tumor markers in drug development; which ones can be used? Some of these tumor markers have been available for three decades. There is already a lot of information on markers such as MRI in joint disease, but we need to be able to look at healing of erosive lesions, for example. Then there are many new biomarkers; how do we develop them? What about pharmacogenomic markers? What about these new "omic" technologies: Will they yield markers, and will those markers undergo the same fate? Will they be published as associations, with larger and larger clinical series, and then some day, will someone use them in monitoring patients? Will they ever get used in drug development? How do medical imaging and all the new functional imaging techniques get incorporated in drug development?

Efforts have been made to turn pharmacogenomics into something that is routinely used in the regulatory world. Many of the drug-metabolizing enzyme polymorphisms are very well understood. So, what we really need, are some commercial tests and drug sponsors who use them in trials. Then we will be able to incorporate them progressively in labels. One of the problems is that the clinical community does not have a good understanding of drug metabolism. We need to find a better way to present these data in a simple and straightforward manner so that the information can easily be used in a prescribing situation. Doing that is going to be just as hard as developing the pharmacogenomics. Pharmacogenomically directed therapies exist. They are being developed especially in oncology. We hope, very soon, that we will be able to screen people for toxicity based on genomic characteristics. The question is, how will we make that a reality and have it actually in use in clinical medicine?

The FDA is going to start a process of discussion about general validation processes, both for biomarkers and surrogate endpoints, not on a theoretical basis but a deep discussion about what needs to be done to make these useful in drug development. We are going to need data from industry; we are probably going to have to do a fair amount of data mining for any one of these markers. Much of the industry has indicated in their submissions to the Critical Path docket, however, that they would be willing to share data. (Perhaps, they would not want it to be identified, but they would be willing to share their data with the FDA for this purpose and allow it to be used to help validate the markers.) We need to discuss the process, the kind of analyses that have to be done, and how to set up collaborations among stakeholders, including NIH, to get these clinical validations done.

34.5.9 Additional Drug Development Challenges

There are other clinical drug development challenges. Evaluating dose–response: Are there better methods for evaluating dose–response? How do we make the regulations make sense so that people feel comfortable doing this work?

Developing quantitative disease models: These models are seen as something that can incorporate all the knowledge into a unifying model. Everything that has been known about the disease so far — response to different therapies, natural history, the progression and fluctuations of biomarkers in the disease, and so forth — could be placed into a model. This provides a tool for people to communicate as well as to test interventions in simulations.

Designing clinical safety biomarkers: Let us take the example of hepatotoxicity. We have had an effort underway for the past 2 years to look at what we know about drug-induced hepatotoxicity. We have come a long way in understanding it. But further effort is needed, for instance, for creation of a disease model of drug-induced hepatotoxicity. What do we know about the physiology of the liver? What do we know about what goes wrong in various instances and examples of drug-induced hepatotoxicity? Can we model further changes and put new biomarkers, such as proteomic or metabolomic markers, into such a model as we gather data? We need to put together an industry–FDA–academic consortium to share the existing knowledge. As companies study potential hepatotoxic drugs, which happens from time to time in diseases where those therapies are really needed, can we conduct studies on clinical trial samples, to search for new markers or to refine our model? This type of effort could quickly advance our knowledge about drug-induced hepatotoxicity.

34.5.10 Rebuilding the Academic Base in Experimental Medicine

The academic base in experimental medicine, in both training and scientific programs, must be rebuilt. Translational research, that is, specific product-related research, alone is not sufficient. A government–academic–industry consortium or multiple consortia may be one model to get some of this research started, but clinical pharmacology needs to be integrated as a key part of the multidisciplinary teams in each of these endeavors.

34.6 Discrete Critical Path Deliverables

The FDA's next step will entail the creation of a Critical Path Challenge list, a series of projects with discrete deliverables, on the basis of what has been submitted to us in the docket and our internal knowledge about the kind of work that needs to be done. The FDA will be able to initiate a few of these, limited by the amount of available resources. We hope that identification and explanation of the projects that are not being done will spur additional efforts to get these projects funded and up and running.

In summary, the scientific infrastructure for drug development must be improved to lower the barrier to new therapies. We cannot keep pursuing the current path. This may sound like a revolutionary statement, but in fact if changes are not made there will not be a surge of new treatments, and the therapies that are developed may not be affordable. Affordability problems are going to become increasingly challenging for the society. The Critical Path initiative seeks to identify concrete projects with specific work products to advance the science and lower the barriers to development of successful medical products. The FDA role is to identify the problems and to collaborate with others on the solutions. Clinical pharmacology will be a pivotal discipline in this effort.

35

Managing R&D Uncertainty and Maximizing the Commercial Potential of Pharmaceutical Compounds Using the Dynamic Modeling Framework

Mark Paich, Corey Peck, Jason Valant, and Kirk Solo

CONTENTS

Part 1: Using Dynamic Modeling to Assess and Maximize Commercial Potential of Developmental Drug Compounds

35.1 Introduction

A vital and yet often neglected part of the drug discovery and development process is the evaluation and maximization of the commercial potential of the developmental compounds that are moving through it. This book addresses a number of issues related to the efficiency of this process and the steps within it. But, in an industry where development costs for compounds are approaching or even surpassing a billion dollars,[1] executing efficiently (doing things right) *must* be combined with operating effectively (doing the right things) so as to have a world-class drug discovery and development system.

For the accumulated costs and resources devoted to the development of a new chemical entity (NCE) or new molecular entity (NME) to make sense financially, the commercial potential of the compound must be evaluated in a rigorous manner. Compounds whose expected financial performance does not warrant these high investment costs must be abandoned or out-licensed as soon as possible so as to direct resources toward more profitable endeavors. By operating effectively, a well-designed drug discovery and development process can focus its efforts to operate efficiently on the compounds that will maximize cash flow to the pharmaceutical firm.

The commercial assessment of NCEs (or biologic entities as well, although for simplicity we will only reference the former in the description that follows) should be an integral part of the new product development process, with three key questions in mind:

1. How can the NCE be positioned to maximize its commercial potential?
   - Testing plans for various phase trials
   - Subsequent market positioning and brand plan strategy
2. Does that maximized potential justify the associated NCE development costs?
3. If not, what options can be pursued to recoup at least some of the accumulated development costs?

Answers to these questions have the potential to be supported by the wide availability of data relevant to pharmaceutical businesses, if only a coherent and consistent framework can be utilized to incorporate it. Extensive records regarding patient and physician behavior are now readily available from claims databases and third-party data providers, which in longitudinal form make it possible to track the dynamics of various patient and prescribing segments over time. Broad competitive intelligence makes product pipelines of industry players transparent to any analyst with an Internet connection. Masses of raw data are useless, however, without an interpretive framework that synthesizes it into information that can generate learning and create the foundation for significant and sustained competitive advantage. And in the ever-changing pharmaceutical world, the dynamic modeling framework is a highly effective organizing tool to capture, analyze, and ultimately leverage the complex dynamics driving industry behavior.

In our consulting practice over the last decade, we have assisted our clients in advancing their brand planning approaches and associated evaluations of commercial potential for developmental compounds toward sophisticated methodologies that have been popular in other industries for years. Dynamic modeling represents a cutting-edge approach for operationally defining the set of inter-relationships that drive the behavior of pharmaceutical marketplaces. This approach is based on established techniques from the world of operations research and system dynamics, and adds value in at least two distinct ways:

1. The dynamic modeling approach creates a means to test the linkages between cause and effect in various pharmaceutical settings. Such simulation-based models can be used to operationally define the expected outcomes of a set of strategic decisions — resulting in better strategic plans and a more complete understanding of the sets of key relationships governing them.
2. Dynamic models create a framework to integrate the knowledge bases that exist within a pharmaceutical firm's functional divisions for the purposes of evaluating an NCE's potential. The approach establishes practical boundaries on the realm of possible outcomes from a commercial assessment, providing a solid foundation for reasonability checks of the financial potential for an NCE. As

such, dynamic models provide an operational view of the future behavior of a disease marketplace and the position of a particular compound within it.

Before investigating the details regarding the dynamic modeling framework, however, it is useful to review the typical process for developing brand plans and commercial evaluations for an NCE, as well as highlight some deficiencies in the existing approaches of strategy development in the pharmaceutical industry.

35.2 Traditional Approach for Brand Planning and Commercial Assessment in Pharmaceutical Industry

The pharmaceutical industry has been a large part of our consulting practice over the last decade, and many of our engagements are headed by marketers in charge of a particular developmental NCE. These marketers often have the title of brand manager for a particular compound and are typically charged with developing a strategic plan for the NCE in late Phase II or early Phase III trials on its way to approval by the Food and Drug Administration (FDA), as seen in Figure 35.1.

Results from the relatively inexpensive Phase I clinical trials force pharmaceutical firms to make critical decisions regarding the future of the NCE and the associated resources allocated to it. Sometime after the beginning with Phase II but often not until Phase III, senior executives will commission a brand plan to assess the technical and commercial potential of the new compound (recognizing, of course, the possibility of compound failure). Brand managers typically assemble a cross-functional team comprising representatives from the firm's marketing, clinical/medical, health outcomes, and forecasting departments to evaluate the financial potential of the NCE (with representatives from regulatory, legal, and even operations rounding out the team line-up). If the new compound is currently in Phase II, the evaluation process might also help decide on a testing plan for the subsequent clinical trials. For example, the strengths and weaknesses of currently available products often suggest a Phase III clinical trail design to demonstrate the superiority of the developmental drug.

The brand plan team is assembled to guide the compound through the development process toward the launch of the product — a complicated endeavor involving numerous

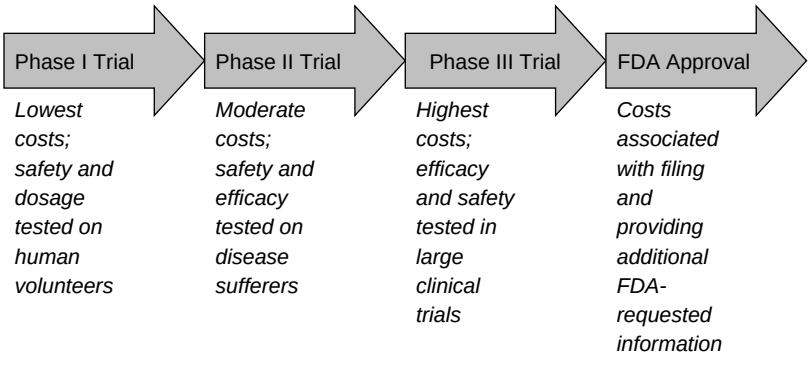


FIGURE 35.1
Clinical trial and FDA approval process for developmental NCE.

milestones and a large investment of human/financial resources. From a marketing perspective, however, a brand plan has two key objectives.

35.2.1 Key Components of a Brand Plan for a New Chemical Entity

1. A set of product strategies for the short-term time horizon (1 to 3 years)
2. An early commercial assessment of the compound's financial potential

Pharmaceutical firms over the last decade have increasingly adopted the practice of cross-functional coordination so as to leverage various types of knowledge within the organization, and marketers often sit at a conceptual hub of company functions when it comes to early assessment of NCE market potential. To develop an effective brand plan, the marketer/brand manager is often the "gravitational force" behind such cross-functional interactions, as depicted in Figure 35.2.

35.2.1.1 Clinical/Medical

Often physicians who previously practiced medicine in a specific disease area, clinical/medical experts, are knowledgeable in the pros and cons of various currently available treatments and have the technical expertise to understand and evaluate the product profile of the NCE, especially as relates to understanding the design and results of clinical trials. Their knowledge of standard practice and corresponding treatment paradigms provides excellent background into how patients are viewed and subsequently treated by physicians.

Clinical/medical experts help place the NCE within a competitive market framework, often by working with the marketer/brand manager to develop strengths, weaknesses, opportunities, and threats (SWOT) analyses.

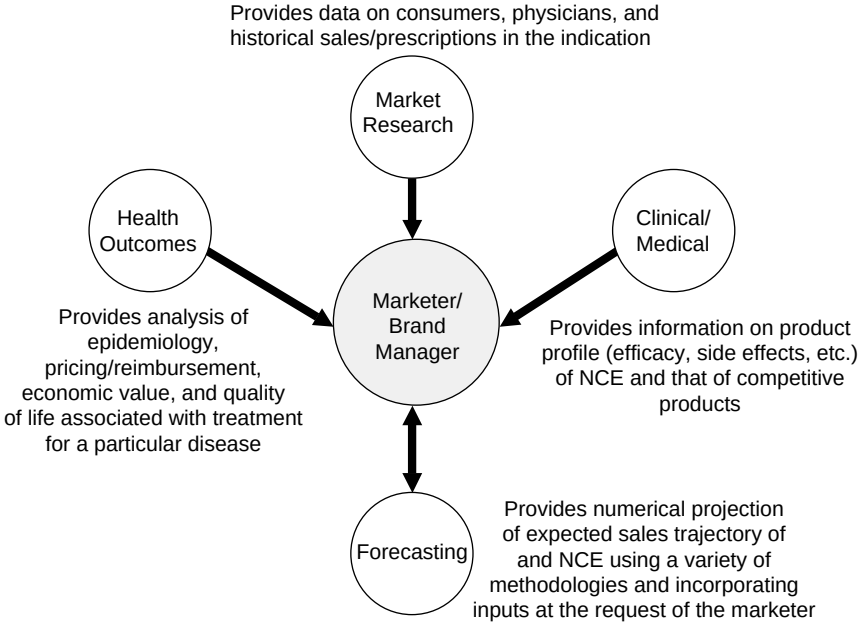


FIGURE 35.2
Interaction of various functional areas in determining a brand plan.

35.2.1.2 Health Outcomes

To assess the impact of a disease on individuals or the economy, health outcomes resources analyze marketplace factors such as epidemiology, the economic benefit of various forms of treatment, pricing and reimbursement, and lifestyle measures of the effect of a disease on quality of life metrics. Epidemiology is an increasingly important part of this analysis and is based on evolving population demographics and changing prevalence, diagnosis, and treatment rates. Health outcomes personnel may provide information how the disease affects specific patient segments at the request of the marketer/brand manager, but their primary contribution is the overall assessment of how a particular disease impacts the health of those suffering from it and lives of those around them.

35.2.1.3 Market Research

Market researchers represent the "voice" of the marketplace and specialize in collecting, analyzing, and communicating information on the indication marketplace and individual existing treatment options within it. Market researchers design primary market research studies to collect data on physician/consumer reactions to NCE product profile, pricing, packaging, etc., often using a conjoint analysis framework. They may also query secondary data sources to obtain specific patient-level data or information on physician behavior. Marketers/brand managers often request specific information from market researchers so as to better understand important facts related to a disease marketplace.

35.2.1.4 Forecasting

As experts in pharmaceutical marketplace analytics and projection methodologies, forecasters can assess the commercial potential of the developmental compound and work closely with the marketer/brand manager to develop a prescription and revenue projection for the NCE. Forecasters are often technically oriented and their role is to provide a numerical forecast for the compound, not focus on the strategic implications of their analysis. However, their interaction with the other members of the brand plan team often clarifies the assumptions associated with the disease marketplace and the place of the NCE within it.

35.2.2 Strategic Outcomes from the Traditional Brand Planning Process

The efforts of the brand plan team result in a comprehensive look at the disease marketplace, the competitive landscape, the pros and cons of various currently available and developmental treatment options, the assessment of the unmet medical need in the market, and other information designed to inform decision-makers about the conditions into which the new compound might be introduced. From a marketing standpoint, however, the key outcome of the brand planning process is the concept of brand positioning — how the compound will be positioned with respect to its target audience, its expected benefits, its key reasons for trial/usage, etc. Brand positioning helps establish a series of product strategies created to leverage the collective knowledge of the disease market and effectively use resources to increase the uptake of the NCE. The strategies are often categorized by areas of target influence, such as patient or physician segmentation, impact on the regulatory environment, effect on pricing/reimbursement, publication strategy, etc. In total, these strategic initiatives are designed to meet some key objectives regarding the performance of the NCE in the marketplace — goals by which the success of the launch of the compound will be measured. The strategies resulting from the brand planning process are then rolled out to the operations personnel in various markets to determine the tactical approach to support them.

35.2.3 Commercial Evaluations from the Traditional Brand Planning Process

At least in part because of the importance of a sound commercial assessment of the NCE, forecasting is usually an interactive and iterative process in which information gleaned from market research, health outcomes, and clinical/medical is given due consideration in developing a financial projection for the NCE. Forecasts in the pharmaceutical world rarely rely on one methodology or a single set of input assumptions. They may be derived from volume analysis, extrapolated from market financial projections, or even calculated on the basis of patient algorithms. Analogs of other drugs in an indication marketplace, statistical formulations, econometric models, and patient-based calculations may all be employed and compared as ways to evaluate the commercial potential of an NCE.

From our experience in the pharmaceutical world, however, these types of forecasting analyses boil down to two sets of aggregate assumptions about an indication marketplace and the position of the NCE within it, as shown in Figure 35.3.

In Figure 35.3, the expected indication prescriptions/sales is usually calculated from an extrapolation of historical data, perhaps modified by epidemiological trends or expectations of advances in the efficacy of various treatment options. Expected share of prescriptions/sales by NCE is generally derived from analysis of product analogs, comparison of the NCE to existing treatment offerings in the marketplace, and team judgment regarding the level and type of marketing support for the compound.

Forecasted values for market size and NCE share are then multiplied to arrive at an expected sales trajectory for the compound (expected NCE prescriptions/sales in Figure 35.3). Note that this projection includes three key concepts:

1. The peak volume/sales figure
2. The speed at which the NCE will reach this maximum revenue potential
3. The shape of the falloff from peak sales

The estimates of an NCE's expected sales trajectory can be analyzed using a Monte Carlo technique, which produces a range of possible outcomes, given uncertainty about the actual magnitudes of various input assumptions. The results of a Monte Carlo analysis

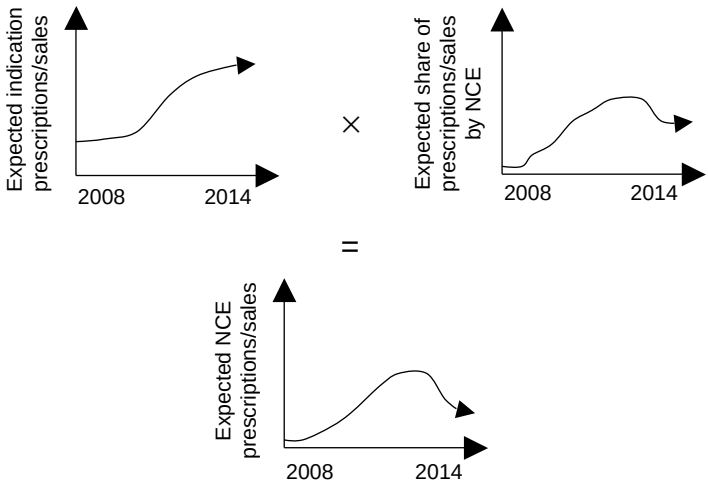


FIGURE 35.3
Basic assumptions in traditional forecasting efforts.

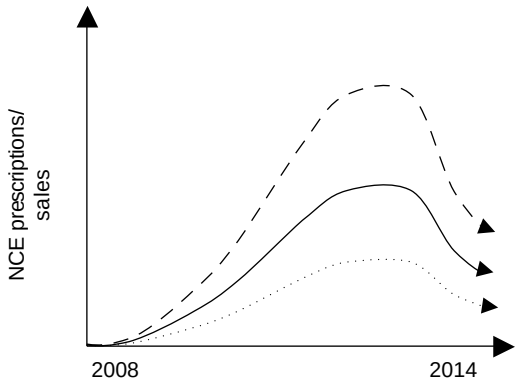


FIGURE 35.4
Examples of stretch, most likely, and worst-case scenario outputs from Monte Carlo analysis.

are often expressed as stretch, most likely, and worst-case scenarios regarding the commercial assessment of the compound, as shown in Figure 35.4.

The result of the typical forecasting process becomes an important part of the brand plan for the NCE, and this forecast is usually presented up the chain of command within the pharmaceutical firm. The early commercial assessment of the NCE often dictates the brand's initial marketing budget, from which a marketer/brand manager can actually develop and execute a strategic plan for the compound. The details of the brand plan are specific to each individual situation, but often leverage the institutional knowledge the marketer/brand manager has gained throughout the course of interactions with various functional areas of expertise within the firm.

35.2.4 Limitations of the Traditional Brand Planning Process

The creation of a brand plan consumes a significant amount of company time and resources — a sound investment if the brand positioning and corresponding strategies are sufficient to meet the stated objectives for the developmental compound. And although the existing brand planning process is the gold standard approach to developing strategic marketing initiatives and robust evaluations of NCE commercial potential, its effectiveness is hindered in four key ways:

1. *Misapplication of product analogs*: Brand plans often include an assessment of the launches of analogous products, related to both the details of successful brand positioning and its corresponding financial performance, but *historical analogs are helpful only in the rare case in which the underlying causal relationships generating market behavior are similar to that of the NCE in question.* The validity of direct analogies to individual products depends on factors such as product similarity, formulary status, the competitive landscape, patient interaction with the health care system, physician choice criteria, and even epidemiology. Without a means to explicitly capture such dynamics and inter-relationships, product analogs are often evaluated on the basis of implicit judgment rather than rigorous analysis.
2. *Failure to leverage the institutional knowledge of cross-functional team members*: While cross-functional coordination has become the standard for the traditional brand planning process, the data/information of these team members are often only

implicitly included in strategic marketing programs or assessment of commercial potential for a compound. Without a common operational framework to incorporate such institutional knowledge, the inputs from various organizational functions are generally not incorporated into the brand planning process in a consistent and explainable fashion.

3. *Inherent limitations of static approaches*: Static tools such as spreadsheets are inadequate to capture the complex set of inter-relationships present in actual pharmaceutical markets. Spreadsheet models often contain detail about a disease marketplace, but are not good at explaining the dynamic marketplace behavior resulting from the interaction between causal factors — the very relationships vital in developing effective marketing strategies. Ignoring these relationships can often result in an evaluation of commercial potential that is fatally flawed.
4. *Maintaining consistent assumptions when evaluating alternative strategic options*: Brand plans often make aggregate assumptions about how various marketing strategies might affect the performance of the compound in the marketplace. Typically, the assumptions behind such strategic programs involve an expectation of the effect on the NCE prescriptions/sales, as seen in Figure 35.5.

The details driving such macro assumptions as well as the expected magnitude and timing of their impact on an NCE's expected uptake are often hidden and unarticulated. Such ideas are often termed as mental models; implicit cause-and-effect relationships that are never explicitly defined or sufficiently explained. Establishing a consistent connection between mental models and their corresponding assumptions is crucial for effective allocation of limited resources and the development of successful marketing strategies yet is difficult to achieve without an operational framework to make such a link explicit and testable.

35.2.5 Lack of Integration between Forecasts and Product Strategies

The practice of brand plan development and its NCE forecast component is designed to leverage knowledge from a cross-functional team, and in a certain sense it does. But because the traditional forecasting process is fundamentally disconnected from the operational realities of the disease marketplace, the marketer/brand manager lacks a tool to quantitatively analyze the effects of the product strategies developed. In other words,

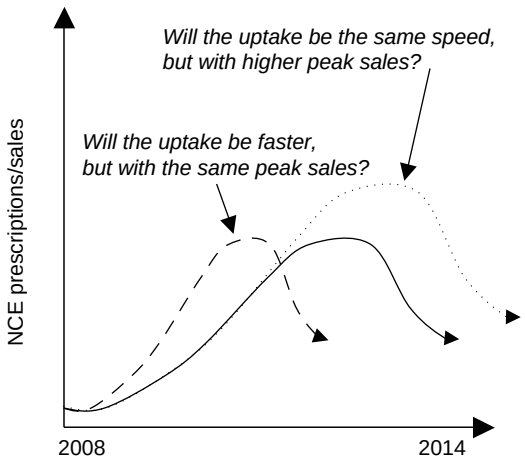


FIGURE 35.5
Macro assumptions of impact of marketing strategies on compound uptake.

there is no established link between brand positioning and the corresponding expected sales trajectory of the compound. Once the NCE is on the market, a marketer can collect data and calculate a return on investment of various marketing programs. In the prelaunch timeframe, however, rigorously testing the effects of possible positioning strategies is impossible without an operational way to evaluate the expected outcomes of those decisions. A brand plan based on such untested mental models is often inefficient with respect to its strategic goals, as it has failed to incorporate the series of causal relationships driving the behavior of the disease marketplace.

35.2.5.1 Risks Associated with Not Integrating Commercial Assessments with Marketing Strategies

1. *Pursuing an NCE of limited potential*: Early commercial assessments often paint a rosy picture, which begins to look tarnished if actual market performance of a compound fails to meet expectations. Allocating resources to products whose revenue potential is limited is costly in two ways. First, the direct costs of such efforts are significant and readily apparent. Second, each dollar spent on a "sinking ship" product means one less dollar available to invest in other, potentially profitable drug development efforts. Forecasting processes such as the traditional one outlined earlier are designed to minimize this risk, but new methodologies to help guide and inform this approach can be beneficial as well.
2. *Abandoning an NCE of high potential*: An understated projection of a compound's commercial value may result in the discontinuation or out-licensing of a product whose revenue stream may turn out to be significant. In the pharmaceutical world, where the revenue streams from successful drugs must fund new development and cover the costs of failed compounds, failure to "ride out a winner" is a recipe for financial disaster.
3. *Failing to strategically position an NCE effectively*: Without a clear understanding of marketplace dynamics, a brand plan may be focusing on low-leverage marketing strategies and consequently fail to maximize a particular compound's potential. Because successful compounds are the exception rather than the rule in pharmaceutical development, maximizing the returns of an NCE as it enters the marketplace is vital to a firm's long-term financial success.

In our consulting experience over the past 10 years, we have seen strategic initiatives and forecasts of commercial potential for a wide variety of diseases, products, and competitive situations. And while each of these two components is a key part of the development process for a compound, they remain fundamentally disconnected in the traditional brand planning approach. This disconnect continues to hamper effective strategic decision-making in the pharmaceutical industry — the costs of which are high and growing. To be truly effective, commercial assessments should at a minimum consider the operational set of causal relationships driving the behavior of a given disease marketplace, as well as incorporate and estimate the effects of various marketing programs on it. *At the level of a brand plan, the expectation of the financial performance of an NCE needs to be operationally tied to the strategies designed to support it.* Without a causal model to link potential marketing actions to expected outcomes, the process of developing, quantifying, and assessing the possible impact of various strategies is done more by gut instinct than explicit and testable analysis. Stated differently, a commercial assessment of an NCE is not leveraged effectively if its forecast methodology ignores the product strategies associated with it in the brand plan.

Our work with pharmaceutical clients focuses on the explicit integration of the commercial assessment and strategic planning processes for effective development of brand plans. The approach utilizes a dynamic modeling framework — a technique and methodology which has been used in other industries for decades to add operational rigor around the structure and expected behavior of particular marketplaces. We have found the dynamic modeling approach to be the single best way to ensure that our clients operate effectively by focusing their efforts on compounds whose expected maximized financial performance merits the associated development investment.

35.2.6 What Is Dynamic Modeling?

> "All theories are wrong, some theories are useful." George Box, but often attributed to W. E. Deming

Useful models are those that are utilized to make better decisions, and people use such models everyday — often without realizing it. Did you take an alternate route to work this morning to avoid traffic? Your revised path was the result of a mental simulation, which predicted you could get to work faster with a "back roads" strategy. Do you invest in a retirement plan? Doing so is likely the product of a mental model which forecasts the long-term positive financial effects of stashing some money away for your golden years. Brushing your teeth, attending college, exercising — all are actions done with the expectation of some derived benefit sometime in the future. In fact, many of the things we do on a daily basis are the results of models — actions or strategies that we expect to have some desired outcome. Though rarely explicitly represented or analyzed, mental models are often the decision-making processes we employ throughout the course of our daily lives.

The situation is not different in the pharmaceutical world, where marketing strategies are developed on the basis of expectations of their effects on drug sales, market penetration, perceived quality, etc. Such strategies are often the product of implicit mental models, and are effective when they have their desired impact. But mental models are difficult to communicate, impossible to analyze or quantify, and hard to prioritize. As such, they often result in suboptimal decisions when it comes to strategy development and brand planning.

Dynamic models represent an operational way to translate mental models from the implicit to the explicit, allowing for a set of specific business questions to be addressed and analyzed. When assessing the commercial potential for an NCE, these strategic questions often center on:

1. Which brand plan strategies maximize the expected commercial performance of an NCE?
   - Should we grow the market, steal share from competitors, or create new markets for the compound?
   - What key market segments are crucial to success of the NCE?
2. How can we better leverage the wealth of data and institutional knowledge about the disease indication, patient behavior, and physician preferences?
   - How can our marketing programs impact the observed behavior of these important stakeholders in the disease marketplace?
   - What vital knowledge of marketplace dynamics is missing from our current information sources?
3. How can we have an integrated tool to link the results of proposed strategies to impact on the compound commercial assessment?

- Is there a way to evaluate how potential marketing strategies will affect the disease marketplace?
- Can a new methodology take into account the causal effects of various marketing initiatives?
- How much does the commercial value of a compound depend on various positioning strategies?

Answering these questions requires a better operational understanding of the fundamental processes driving the expected evolution of an indication marketplace — a process that utilizes the structure/behavior paradigm.

35.2.7 The Link between Structure and Behavior

The defining aspect of the dynamic modeling approach is its focus on the link between the *structure* of a system and its resulting *behavior*. Dynamic models go beyond statistical correlations or regression algorithms to address the fundamental processes and interrelationships that define system results. Such understanding is essential to effective strategic decision-making. To quote Barry Richmond, a visionary in this arena, "By thinking in terms of how a process or system really works (i.e., its "physics"), we have a much better chance of understanding how to make it work better!"[2]

The dynamic modeling framework forces an understanding of the structure of an indication marketplace to formulate a better, more effective set of strategies for maximizing the commercial potential of an NCE. Knowing the uptake trajectories of previously released compounds in the indication or isolated data on epidemiology, treatment paradigms, physician attitudes, and competitive products is not enough. Rather, an integrated approach that combines the effects of these operational factors can not only explain the sales trajectories for product analogs but also provide a means to quantify such key strategic questions as:

1. How many patients are newly diagnosed each year?
2. How many patients are switching treatments?
3. What drugs are they switching from and to?
4. How are changing patient demographics and epidemiology likely to affect the key patient dynamics?
5. What stages do physicians progress through in their acceptance of a new compound?
6. How do marketing levers affect the speed at which that diffusion takes place?
7. What are the currently available treatment options in the indication?
8. How do these therapies compare to one another and the NCE?

Answering questions such as these depends on a clear understanding of marketplace dynamics and the structure/behavior paradigm, resulting in effective and successful brand plans and corresponding marketing strategies. Structure and behavior links are pervasive in the pharmaceutical industry, and the dynamic modeling approach can serve to frame and analyze the sets of inter-relationships that ultimately drive marketplace evolution.

In our consulting practice we have developed a structure/behavior framework that provides a useful way to examine an indication marketplace, particularly when assessing commercial potential of a developmental compound, as shown in Figure 35.6.

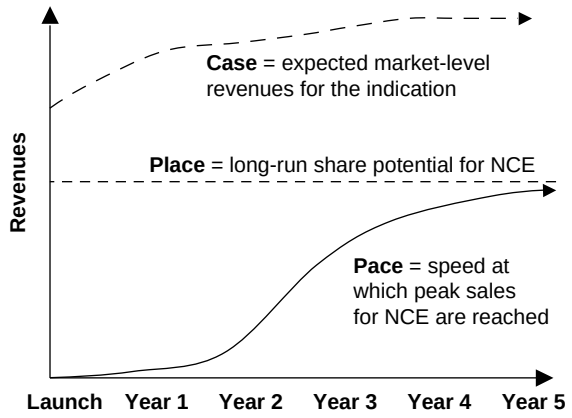


FIGURE 35.6
Case, Place, and Pace framework for an NCE in a given indication.

1. The Case for a new compound is the set of aggregate, macroeconomic market conditions of the indication evaluated. Information on epidemiology, changing demographics, treatment paradigms, historical market revenue trends, and existing competitors all define the case for an NCE considering entry into a new marketplace. Case issues are often summarized as revenue forecasts for an indication, although by now readers should readily recognize the underlying operational dynamic assumptions behind such projections. The concept of case is vital when evaluating the feasibility of continued support of a developmental compound, as it determines the aggregate conditions into which such a product introduction will take place. The strategic influence of individual pharmaceutical firms on case issues must be carefully evaluated, however, as by definition strategies designed to change the case redefine the market environment for all players in it — resulting in a larger pie for collaborators and competitors alike.
2. The Place for a new compound defines how it compares to the existing and future competition in terms of efficacy, safety, side effects, etc. In markets exhibiting specific treatment algorithms, place may also include where the compound is expected to fall within a defined lines of therapy progression. The place for an NCE is in some ways determined by its FDA-approved label, which can limit a compound's prescribed usage to certain segments of the patient population on the basis of metrics such as disease severity or failure of existing treatment options to control the disease. In economic lingo, the place of a compound in an indication determines its long-run market share potential — often referred to as peak sales in the pharmaceutical world. This upper bound of commercial potential for a compound is under the influence of marketing efforts, however, as increasing a compound's perceived treatment attractiveness or pursuing a labeling strategy to get the compound earlier in a line of therapy progression can impact where that long-run share ceiling is observed. The specific design of clinical trials may also factor into the equation, as the outcomes of such trials are often key determinants in labeling and perception of NCE utility in the existing marketplace.
3. The Pace for the new compound relates to the speed at which it can reach peak share potential — the shape of the expected sales trajectory. While case and place are generally functions of product attributes, disease epidemiology, and other aggregate conditions, the pace for an NCE is more commonly under the influence of various strategic marketing decisions of the part of individual pharmaceutical

firms. Effective brand planning leverages the observed dynamics in the marketplace to effectively increase the pace at which an NCE is accepted in the marketplace. The resulting expected sales trajectory, commonly referred to as product uptake, is an important factor in the commercial assessment of the compound using financial metrics such as net present value.

The structure/behavior framework of Case, Place, and Pace for a developmental compound is driven by a combination of dynamics, which ultimately determine marketplace evolution.

35.2.8 Three Key Sectors of the Dynamic Modeling of Pharmaceutical Marketplaces

We have developed what we call the standard template dynamic model that answers some common strategic questions for an NCE by combining aspects of patient flow dynamics, doctor adoption of pharmaceutical products, and the perceived treatment attractiveness of the treatment options within a disease marketplace. These components are populated with data from epidemiology and a variety of physician and patient databases to ensure a robust representation of actual market dynamics (see Figure 35.7). Merging these three structural pieces into a dynamic model determines what drugs patients receive, how long they remain on treatment, how they transition between therapies, how rapidly doctors accept new forms of treatment, and the evaluation of available treatments in the marketplace. In the standard template, these three key components come together to form an integrated model that captures and quantifies the important dynamics that operationally determine how pharmaceutical markets behave over time. Leveraging this knowledge in an operational simulation environment provides the means to test the effects of various strategic options.

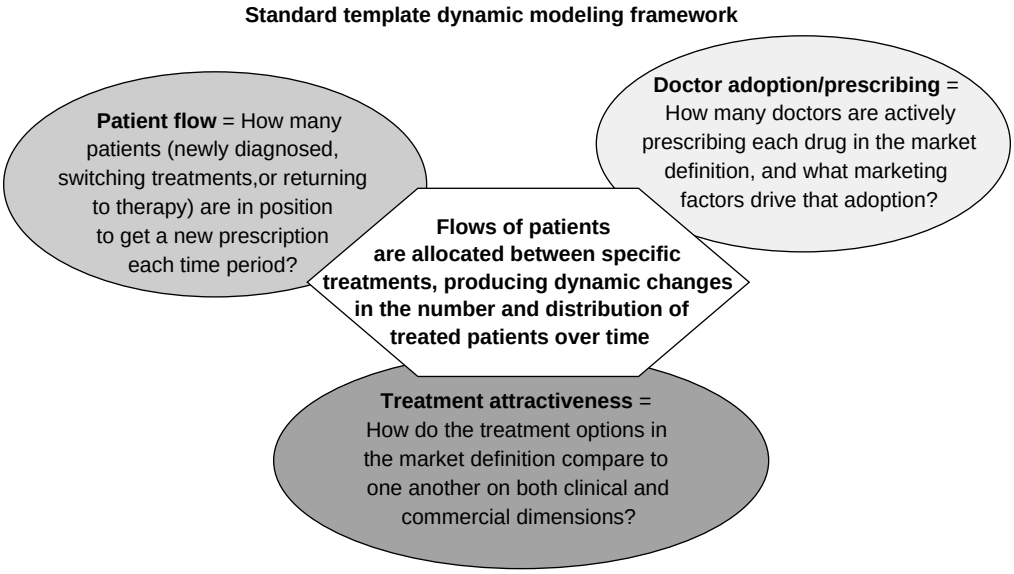


FIGURE 35.7
The three sectors of the standard template dynamic modeling framework.

35.2.8.1 Patient Flow

Product-based businesses rely on giving customers the opportunity to use a particular offering, and a number of different frameworks exist for analysis of consumer goods markets.[3] The operating principle for pharmaceutical products is getting patients to try and continue to use an individual drug therapy. Pharmaceutical firms often collect or purchase data on the number of prescriptions written or filled, as prescriptions can be translated into a revenue estimate for a particular drug in a given indication. Yet, rarely does the effort ask the fundamental question of the driving forces behind prescription generation.

Instead, early commercial assessments of NCEs often employ statistical analysis to produce projections of the numbers of prescriptions in an indication marketplace through time. But from a causal standpoint, forecast methodologies do not address the underlying factors that determine how prescriptions are created in the marketplace. Operationally, prescriptions are generated by patients as they

1. Initiate prescription treatment for the first time
2. Switch from one treatment to another
3. Return to treatment after having been not treated for a certain time period
4. Refill their existing prescriptions periodically

These treatment opportunities represent the flow of patients in a given indication over a specific period of time. There is a famous saying that the education at MIT is like trying to take a sip from a fire hose, which in the end is better than taking a large gulp from a smaller information stream at a less prestigious university. The same principle applies in a pharmaceutical context, but relates to the flow of patients moving within a marketplace and the number of those patients captured by individual compounds. *Operationally, the volume of patient flows and their allocation between various treatment options in a marketplace fundamentally determine how the magnitude and associated shares of patients/prescriptions/revenues will change over time.* From a strategic standpoint, the questions become, "How big are the relevant patient flows, how many patients might a particular NCE capture, and what strategies would help capture them?" Only by answering these questions can an operational understanding of NCE commercial potential be determined. The standard template model defines the magnitude of these treatment dynamics under the heading of Patient Flow, and establishing such an operational structure for prescription generation and the resulting analysis is extremely important for effective strategic planning and commercial assessment.

35.2.8.1.1 Case Study: Operationalizing Expected Sales Trajectory

A pharmaceutical firm is launching an NCE into a disease marketplace of 1,000,000 currently treated patients, and wants to know how many patients might be captured during a critical 3-year launch period. Analogs and an early commercial assessment have indicated that a 20% patient share is possible at 3 years postlaunch — somehow the firm must capture (and keep!) 200,000 patients over the next 3 years. Figure 35.8 depicts the key operational question behind such an expected sales trajectory.

Traditional projection methodologies typically say *nothing* about the patient dynamics that ultimately drive the type of uptake curve shown in Figure 35.8. Operationally, this accumulation of patients results from capturing and keeping a combination of patients who are switching treatment, reinitiating treatment, and others who are completely new to treatment. Most commercial assessment efforts imply patient numbers at given time points, such as 200,000 patients at the 3-year mark, but are silent as to the operational assumptions behind such projections. The more important question from a strategic standpoint is

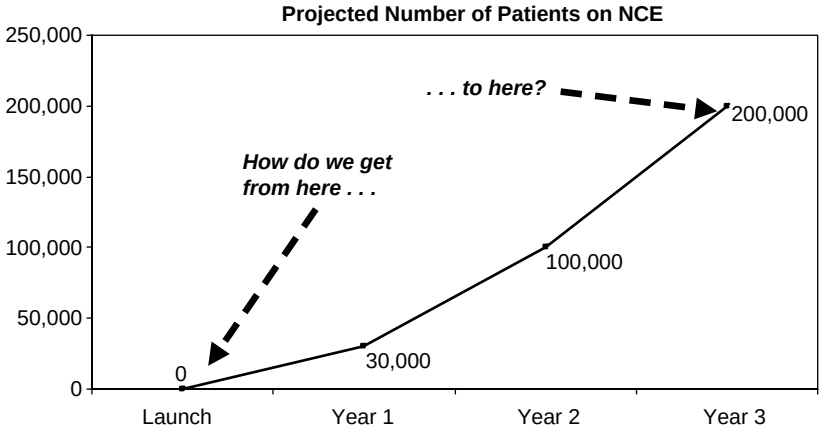


FIGURE 35.8
Operational understanding of patient projections.

TABLE 35.1
Operationalized Mathematics behind Example Marketplace Patient Flow Dynamics

Patient Flow	Total Patients/Year	NCE Captures (%)	For How Many Years?	Total Patients on NCE
New to treatment	10,000	100	3	30,000
Switching treatment	20,000	100	3	60,000
Reinitiating treatment	5,000	100	3	15,000
Total	**35,000**			**105,000**

whether such an expected trajectory is reasonable, given observed marketplace dynamics surrounding the flow of available patient opportunities.

Suppose a series of market research and data analysis projects indicated that only 10,000 patients are new to treatment every year, another 20,000 switch therapies, and only 5000 patients return to treatment over a year's time. As Table 35.1 indicates, even assuming the new compound captures every single one of those patients and keeps them indefinitely, the original estimate of 200,000 patients at 3 years postlaunch begins to look incredibly optimistic.

Note that from an operational standpoint, Table 35.1 suggests that the expectation of 200,000 patients is simply not achievable. Marketing strategies could be designed to increase the volume of patient flows in the marketplace, but even back-of-the-envelope numbers such as those shown in Table 35.1 begin to provide some boundaries for such assumptions. Note that marketing campaigns would have to double the respective patient flows for the entire 3-year period to even have the possibility of hitting the 200,000 patient mark, and even that assumes a 100% capture rate. This type of analysis, on the basis of quantifiable and operational dynamics regarding patient opportunities, allows the reasonability of assumptions to be examined. As importantly, these types of questions are rarely even asked without the operational perspective of the dynamic modeling process.

Dynamic models operationally define and quantify the flows of patients, which represent opportunities in a given marketplace. In general, the patient flow component recreates the structure and magnitude of patient movements within a simulation environment representing an indication marketplace. This type of operational examination of market dynamics becomes a useful commercial evaluation tool. In addition, it can help inform and test key marketing strategies considered for an NCE's brand plan designed to maximize that potential.

The differential behavior of various patient segments can also be included in the patient flow sector as well. Epidemiological concepts such as incidence, prevalence, diagnosis, and treatment rates often vary across segmented dimensions such as age, gender, and disease severity. The patient flow structure is fully flexible in terms of segmentation, allowing for a more accurate representation of differential dynamics of various patient groups.

35.2.8.2 Doctor Adoption

In many pharmaceutical drug markets, historical data can be analyzed to determine how quickly drugs in a particular indication were adopted by prescribing physicians. Prospectively, market researchers often collect data on awareness of drugs that have yet to be launched and conduct surveys on doctors' expectations regarding their future prescribing patterns. This information, while useful, does not explicitly capture the mechanisms at play in both past and future adoption of pharmacological treatments by doctors, nor does it indicate how such dynamics interact with other parts of the marketplace to affect overall market performance. The dynamic modeling framework can be applied to issues surrounding doctor adoption to:

1. Understand the operational physics driving adoption of pharmaceutical products
2. Determine the importance of certain marketing levers in driving that adoption
3. Relate the aspects of physician adoption to components of patient dynamics and various treatments in a competitive landscape
4. Develop strategic forecasts on the basis of an integrated view of marketplace dynamics

Dynamic modeling utilizes a physician adoption structure that is supported by long-standing statistical approaches to product diffusion and can replicate the historical behavior of various drug analogs for a particular indication. This analysis establishes a visual, operational representation of the doctor adoption process and creates a powerful simulation tool to test the effectiveness of various physician marketing strategies. Published marketing literature can provide estimates regarding the effectiveness of various marketing expenditures specific to the pharmaceutical industry. These diffusion factors are easily translated into the doctor adoption sector of the standard template framework, quantifying the effect of various physician marketing strategies on the overall marketplace both now and in the future.

Doctor adoption frameworks generally begin with aggregate analyses of physician behavior, which can then be disaggregated into more detailed doctor segments along lines of specialty, adoption propensity, and even disease-specific characteristics. For example, general practitioners (GPs) may be accounted for separately from specialists in an indication, including word-of-mouth or referral patterns between these two groups of potential treating physicians. Data analysis or attitudinal studies may reveal that certain doctors tend to adopt new products faster than their peers, suggesting a further model disaggregation according to adoption propensity. Dynamic models often include adoption of drugs specific to certain patient segments — a feature which is especially useful for indications in which a new treatment may be readily accepted by physicians for their most severe patients, for example, but not for patients suffering from only mild symptoms. The ability to incorporate and test the adoption behavior of different physician segments into a dynamic model provides a more realistic and detailed picture of how effective marketing strategies related to physician acceptance of products can best be developed.

35.2.8.3 Treatment Attractiveness

Consumers on a daily basis make purchasing decisions on the basis of perceptions of how products stack up against one another. Many factors may affect perceived product attractiveness: price, function, look, style, status, etc. And while products vary across indication and consumers are not identical in their choice criteria, understanding the process through which purchasing decisions are made is vital when analyzing consumer markets.

A similar evaluation process takes place when physicians make prescribing decisions for patients with a particular disease. A set of product criteria is used to determine which drug is most appropriate to prescribe in a given situation. Dynamic modeling allows the concept of therapy attractiveness to be applied to treatment options in a particular indication, allowing the brand plan team to determine the relative utility of various therapy regimens in an indication. More importantly, the dynamic modeling process can help determine the degree to which these differences affect prescribing decisions in the marketplace.

Treatment attractiveness is an overall metric of utility that can be a function of product attributes such as safety, side effects, price, etc. The evaluation of attractiveness can also be tied to specific patient segments or physician specialties to more closely replicate differential evaluations in a complex marketplace, providing an additional level of analysis and insight into the standard template.

35.2.8.4 Structural Example Standard Template Dynamic Model

Dynamic models are usually custom built to meet specifics of a given disease marketplace and the needs of a particular client team. However, a basic framework of the standard template is often a useful first pass in integrating the sector dynamics at play in many chronic indications, as shown in Figure 35.9.

The iconography of Figure 35.9 is a vital part of the dynamic modeling approach, as it makes a clear distinction between places where things accumulate (called stocks), the movement of items between accumulations (called flows), and the information, data parameters, and decision rules that govern those movements. For example, the box called undiagnosed patients is a stock concept, as it represents a place where a number of patients can be counted and quantified at a single time point. Flows of *true incidence* and *diagnosed incidence* bring patients to or take patients from this stock over time. (Flows of mortality are not shown here for the sake of simplicity, but are included in most dynamic models for a full and integrated representation of disease epidemiology.) This flow of diagnosed incidence can be further broken down into an allocation of patients to a range of treatment options (denoted by the stacked stock called currently treated patients), which is itself a function of factors such as the number of prescribers for various therapy regimens and the associated *treatment attractiveness* of those regimens according to evaluations of their *efficacy*, *safety*, and *price*. Dynamic models can be expanded from the basic framework shown in Figure 35.9 to create much more detailed and representative causal relationships present in various therapeutic areas. In this manner, standard template dynamic models can be populated with data estimates, correlated to historical behavior of a given marketplace, and set in motion via simulation to see the likely evolution of market dynamics. Not only is this process a much more robust way to evaluate the commercial potential for a compound given changing marketplace conditions, but it also provides an operational tool to test various positioning and marketing strategies.

35.2.9 Overview of Dynamic Modeling Advantages

Standard template dynamic models represent a departure from traditional approaches, but have a number of distinct advantages that make their application a vital part of effective brand planning and commercial assessment.

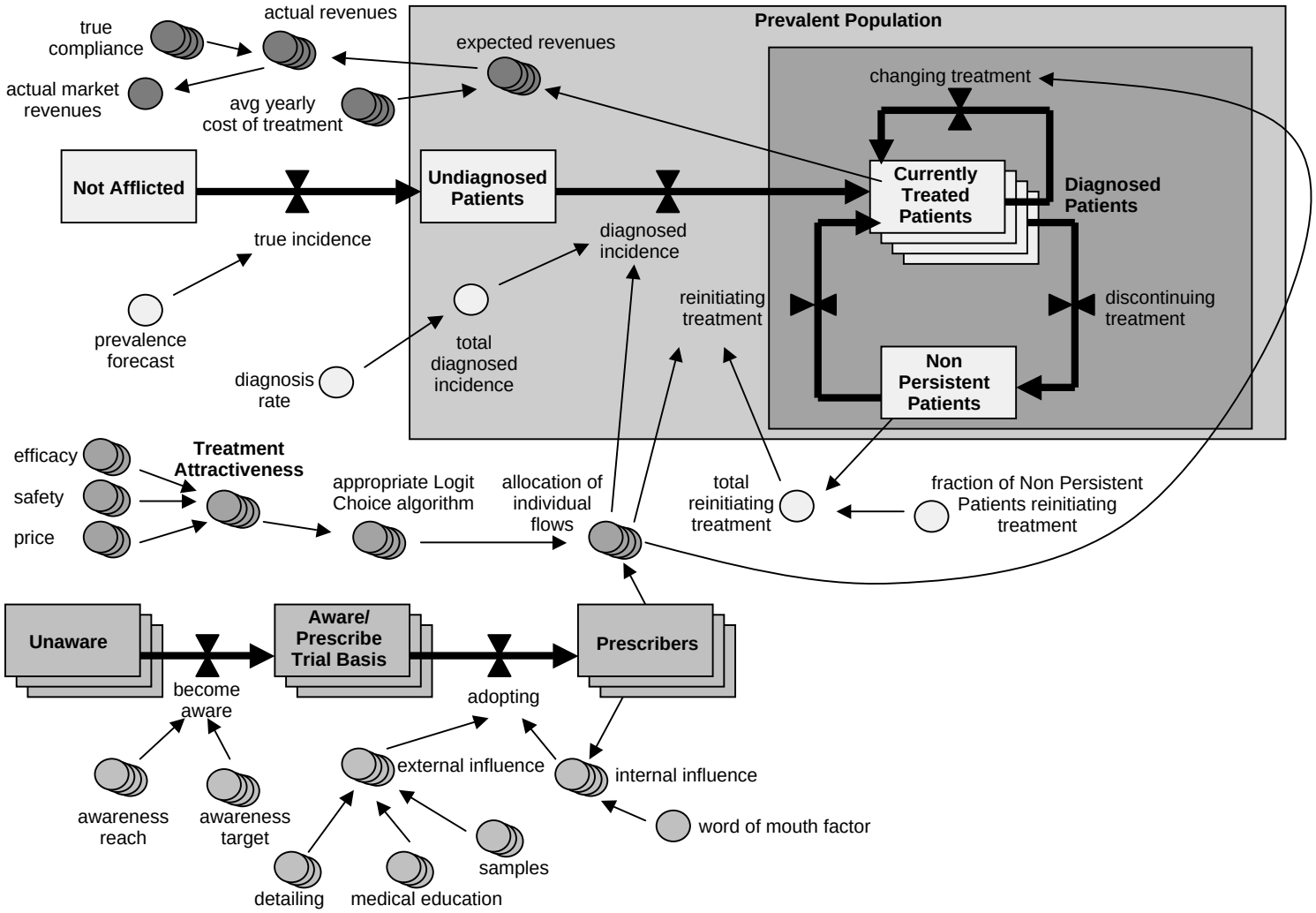


FIGURE 35.9
Standard template dynamic model for chronic indications.

35.2.9.1 Robust Commercial Evaluation Methodology

Many pharmaceutical companies generate an early commercial assessment of an NCE for planning and strategy development, and traditional forecasts may estimate the number of patients, volume of prescriptions, or the resulting level of sales for a particular new compound in each year after its launch. These projections may not be realistic in the current market environment, however, and dynamic modeling provides a much more operational way to generate such evaluations.

Because the standard template takes inputs regarding the comparative utility of treatment options and calculates resulting patient shares and associated product revenues over time, it can be used to determine how good a new compound must be to attain a certain market share target. These types of analyses often prove useful in the early stages of drug development when "go, no-go" decisions are made.

The epidemiology factors of a disease are another area where dynamic models provide important insights and clarity for marketing teams. The explicit structure of disease progression and the associated epidemiology metrics are rarely included in traditional forecasts. However, the dynamic modeling framework clearly accounts for these factors by showing the relative magnitude of patient dynamics as they apply to the epidemiology of a disease. Metrics such as true incidence, diagnosed incidence, point- and lifetime prevalence, and recovery/relapse rates are all typical outputs of the standard template. Identifying the magnitudes of and relationships between these epidemiology concepts provides a useful perspective on a disease and its associated market.

35.2.9.1.1 *Case Study: Misguided Strategy Based on Inaccurate Understanding of Epidemiology*

For example, one of our clients was formulating a strategy of market penetration on the basis of capturing newly diagnosed patients. Their spreadsheet model was a "bottom-up" approach on the basis of projected number of doctor visits, the number of visits at which the medical condition could be identified, likelihood of a prescription, etc. The detailed analysis indicated a strategic opportunity to tap into large flows of diagnosed incidence, and the resulting forecasts showed a very quick uptake of the NCE on the basis of this strategy. However, the static approach did not account for the relationship between incidence and prevalence captured in an integrated epidemiology framework. A very basic dynamic model, populated with the client's derived epidemiology data, showed a huge inconsistency in incidence/prevalence assumptions. The incidence rates calculated by the bottom-up approach implied that the prevalence of the disease would nearly double in only 3 years. This quick sanity check using dynamic modeling principles allowed the client team to truly understand the epidemiology of the disease and abandon a doomed strategy of pursuing newly diagnosed patients — a strategy based on a flawed understanding of key epidemiology concepts.

35.2.9.2 Compliance/Persistency

Compliance and persistency are typically combined in traditional analysis of pharmaceutical markets into one metric to account for patients who do not take all of their prescribed medications. We have found, however, an important psychological and dynamic difference between these two concepts. In the patient flow sector of the standard template, compliance is a continuous metric to account for patients who continue to fill their prescriptions but do not take 100% of their prescribed dosage. Noncompliant patients are a continuing source of revenue, though at a reduced rate because of the corresponding extension of time between prescription refills. In dynamic modeling parlance, persistency relates to how long patients stay on any form of treatment before discontinuing prescription medication entirely.

Nonpersistent patients are no longer revenue-generating as they have ceased to be treated by any prescription therapy. These concepts are clearly depicted in Figure 35.8, and this clean separation allows the patient flow framework to operationally address some important questions regarding patient dynamics and resulting effects:

1. How do assumptions regarding compliance affect overall revenue projections?
2. What are the market-level persistency metrics?
   - How long do patients stay on treatment?
   - Is persistency different for different treatment options?
   - How many patients are nonpersistent at a point in time?
3. What do persistency metrics in the indication suggest regarding strategic planning?

Accounting for compliance and persistence separately in the standard template allows for a better understanding of the patient dynamics around these key issues, and hence a more representative evaluation of a compound's commercial potential on the basis of these distinct concepts.

35.2.9.3 Integration of Short- and Long-Term Forecasts

Many pharmaceutical firms have short- and long-term forecasts created by separate teams using different and even conflicting methodologies and assumptions. Reconciling these approaches is often impossible, resulting in confusion and inconsistency at various levels of the organization. Most pharmaceutical teams know that the long-term performance of a compound is simply the result of a series of short-term performance metrics — just as the flight path of trans-Atlantic airplane is a function not only of the plane's takeoff but also subsequent changes in altitude, speed, and direction. But often these teams have no way to compare short-term data patterns with the longer term dynamics, which generally dominate the commercial evaluation of a new compound.

In contrast, the dynamic modeling framework is not dependent on a specific time horizon, and as a simulation-based approach is able to calculate outputs in any useful time frame. A dynamic model can be programmed to show monthly dynamics, which are especially important at the very early stages of a drug launch. These monthly metrics can be aggregated into quarterly or yearly outputs, which are more appropriate for long-term forecasting or strategy development. Having one tool using a single dynamic modeling approach and populated with identical data ensures internal consistency between short- and long-term commercial evaluation efforts.

35.2.9.4 Determining Strategic Leverage

Operational representations and quantifications of marketplace dynamics provide a means to rigorously test potential positioning and marketing strategies to find points of leverage. A dynamic model can establish a set of base case outputs, on the basis of a defined structure and associated data inputs, which can then be tested to determine their strength of influence. Brand plan teams are often eager to understand the relative impact of these variables to formulate effective marketing strategies.

Sensitivity analysis can be performed in a number of ways, and the flexibility simulation allows for a wide range of dynamic sensitivity analyses to be performed and analyzed:

1. *Basic*: changing the magnitude of each model input by some constant.
2. *Magnitude*: testing the differential effect of changing an input variable by differing amounts.
3. *Timing*: varying the time at which a change in an input variable takes place.

4. *Combinations*: checking combinations of variables and strategies.
5. *Optimization*: running multiple simulations and testing all combinations of strategies to find the optimal solution.

With an integrated dynamic model, sensitivity analyses shows *how the entire indication marketplace will be affected over time* — something not easily or commonly done in other analytic approaches. Sensitivity analysis and optimization represent a useful way to quantify the possible effects of strategies created during the brand plan process, as well as determine a range of possible outcomes for the commercial value of a compound.

35.2.9.5 Decision Analysis

Marketing teams commonly employ a strategy table comparing various strategic options designed to reach their operating objectives or target forecasts. This list of possible strategies is usually organized by potential marketing initiatives and their expected financial return for the compound. Brand plan teams often spend days debating the pros and cons of various strategies, the reasonableness of their associated financial metrics, and the pitfalls in their tactical implementation. The results of such evaluations are a small group of strategies in the form of a brand plan designed to meet the financial goals of the compound in the short- and long-term.

The drawback of decision analysis on the basis of a common strategy table is that *the assumptions behind various strategies that ultimately drive associated financial performance of the NCE are rarely explicit and hence not subject to scrutiny or debate*. Decision analysis sessions can sometimes devolve into turf wars in which political clout and strength of personality substitute for rigorous analysis and open discussion/evaluation of strategic options. In our experience, marketing teams frequently agree to the strategies behind a brand plan without fully understanding the true nature of the assumptions behind it. In the parlance of dynamic modeling, these situations arise from a failure to understand the link between structure (the components, relationships, and assumptions driving a particular strategy) and behavior (the expected result of that strategy).

The dynamic modeling methodology offers an intuitive and visual framework for establishing an operational link between possible strategic actions and their expected outcomes. Utilizing such a dynamic model allows marketing teams to undergo the process of decision analysis in a way that *explicitly* captures the knowledge of various team members and company functions. The resulting simulation tool provides the means to rigorously compare the results of various strategic options. Doing so elevates the process of evaluating a set of strategy table possibilities from argument to assessment, as team members are able to debate not just a set of strategic expectations but also the detailed assumptions driving each. In this manner, the dynamic modeling process leverages team knowledge and information more effectively, resulting in better strategic decision-making in the world of brand planning.

35.2.9.6 Limiting Physics

Traditional early commercial assessments of NCEs often fail to account for the dynamic complexity that ultimately drives behavior in an indication market. The dynamic modeling process, however, captures a series of operational dynamics driving marketplace evolution. This analysis often shows behavioral limits imposed by the dynamic structure — what we like to call the physics of the system. Limiting physics dictates practical boundaries for how an indication marketplace can be expected to evolve, and forecasts that exceed these constraints or the marketing plans built from them are in direct conflict with the underlying dynamics at play. Failure to account for or recognize such limiting physics can result in flawed or even fatal strategies.

Dynamic modeling can help marketing teams understand the operational physics driving marketplace behavior and which strategies can take best advantage of those dynamics. By focusing effort on areas of high leverage as determined from a dynamic modeling exercise, brand planners can better use their resources and pursue more effective strategic opportunities to maximize the commercial performance of a developmental compound.

35.2.9.7 First-to-Market and "Me-Too" Drugs

One of the most powerful aspects of the dynamic modeling methodology is the incorporation of time and the associated delays involved in launching a new compound. Because dynamic models are simulation-based, they have the flexibility to "step forward" in time according to any specified interval. Explicitly capturing the time dimension of key market dynamics allows pharmaceutical teams to test the implications of getting to market first versus coming in later as a me-too drug, allowing the importance of entry position into the marketplace can subsequently be evaluated and analyzed. These types of metrics are especially important when investigating the cost/benefit tradeoffs of increased investment designed to accelerate the development time of a particular NCE.

35.2.9.8 The Role of Data in Dynamic Modeling

Although pharmaceutical companies spend millions of dollars on collecting and storing data of various types, *the real competitive advantage lies in properly using and analyzing this information effectively.* As brand plan teams develop a dynamic model structure, the resulting picture of the marketplace helps focus efforts on the most relevant and useful pieces of information necessary both to complete the model and to answer some important strategic questions. In dynamic modeling projects, data collection becomes an integral part of the overall decision-making process that exists in concert with, rather than isolation from, overall strategic objectives.

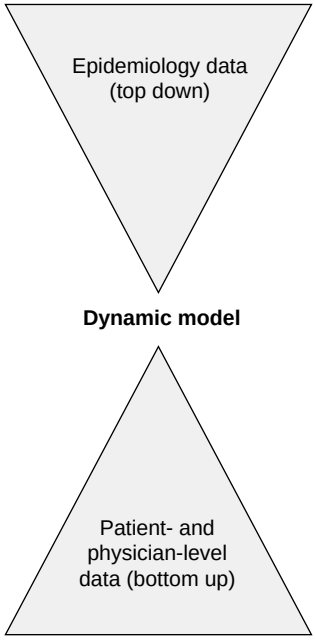


FIGURE 35.10
The role of dynamic modeling in integrating various data sources.

As shown in Figure 35.10, dynamic models provide an integrated, operational platform to combine aggregate, top–down data such as epidemiology with detailed, bottom-up information about patient and physician behavior that is available from a variety of third-party vendors. The dynamic modeling methodology ensures internal consistency of data by integrating it into a single simulation environment that checks whether individual parameters make sense relative to one another. This process gives decision-makers the confidence to utilize and leverage the wealth of information collected from various functional departments, as well as focus the efforts of scarce time and resources when determining the crucial data for new compound commercial evaluation and strategy formulation. Data represent the engine of a dynamic model, allowing defined structure to be simulated to produce expected behavior. As such, it represents a key and sustainable competitive advantage when evaluated and utilized effectively.

35.2.9 Summary

Dynamic modeling represents a cutting-edge approach to evaluating and maximizing commercial potential of newly developed compounds in the pharmaceutical industry. By establishing explicit and operational structure that represents the physics of a system, dynamic models go beyond basic statistical correlations or numerical exercises. In addition, they create a simulation environment in which causal relationships allow for the impact of various marketing and positioning strategies on an NCE's commercial evaluation to be tested. Dynamic modeling is an approach and methodology that can help the drug discovery and development process to be more effective by focusing effort on compounds whose maximized commercial potential warrants their associated development costs.

35.3 Part 1 Recap

1. Rigorous evaluations of the commercial potential of new compounds are a vital part of the drug discovery and development process.
2. A number of inadequate analytic methodologies are currently used to evaluate pharmaceutical markets and develop strategies therein.
3. Strategic decisions are often made based on implicit mental models that are neither testable nor well-communicated.
4. The contributions of cross functional inputs to the existing strategic planning process are not efficiently utilized.
5. The risks associated with ineffective commercial assessments are huge in the pharmaceutical world.
6. The need for an integrated approach of tying commercial assessment of compounds to the marketing strategies designed to support them is met with the Dynamic Modeling approach.
7. All models are wrong by definition, but dynamic models have the advantage of incorporating operational physics in explaining marketplace behavior.
8. A Standard Template dynamic model incorporates aspects of Patient Flow dynamics, Doctor Adoption of newly-release pharmaceuticals, and Treatment Attractiveness metrics of indication-specific therapies.
9. Dynamic models offer several advantages to strategic planning and commercial assessment that focus on an operational understanding of marketplace dynamics and a scenario tool creating cause-and-effect linkages between potential marketing initiatives and their expected outcomes.

Part 2: Risk Assessment and Addressing Uncertainty in Pharmaceutical New Drug Discovery and Development

35.4 Introduction

Portfolio management concepts have been an accepted component of the financial services realm for decades, and the need to diversify risk continues to be an effective organizing principle for many financial instruments and analytic techniques. Risk assessment is also a key component in the pharmaceutical world, where as few as 1 in 5000 laboratory-tested compounds ever makes it to the market.[4] As such, the pipeline portfolio is a critical indicator of the future of a pharmaceutical firm, and the process of analyzing and managing it effectively is of vital strategic importance.

Pharmaceutical pipeline portfolios, however, have industry-specific nuances, which present some significant challenges. Traditional financial portfolio analysis often fails to address these subtle distinctions, while pharmaceutical domain experts often lack the financial portfolio background required to adapt those approaches to pipeline issues. The mismatch results in decision analysis on the basis of a suboptimal set of evaluation tools. This type of inadequate analytic methodology has huge financial implications in the pharmaceutical world, where yearly research and development costs are now over $50 billion globally.[5]

The dynamic modeling methodology can be applied to pipeline portfolio issues in a way that integrates accepted financial portfolio concepts with the nuances specific to the development of pharmaceutical compounds. Such models serve as a litmus test for the future health of the enterprise by providing a useful and accessible tool for analyzing the possible evolutions of the product pipeline and understanding the implications of various strategic options for managing it. Many of our clients use dynamic models in this realm because of their ease of implementation, their sophistication of analytic techniques, and their flexibility to adapt to industry- or company-specific distinctions in the pharmaceutical industry.

35.5 Establishing a Dynamic Model Simulation Framework for the Drug Development Process

Dynamic modeling is an approach that first establishes an explicit structure which depicts the operational nature of a given system or process, then populates that structure with specific data related to it. When a populated dynamic model is "set in motion" via computer simulation, a range of expected outcomes can be determined and analyzed.

A brief introduction to the principles of dynamic modeling often helps describe the methodology before examining how these principles can be applied to the drug discovery and development process. Dynamic models first use a pictorial language to describe the fundamental processes operating within any system. This language makes a clear distinction between the places where items accumulated, the rate at which items move between those accumulations, and the factors or decisions that govern the rate at which those movements take place. In dynamic modeling parlance, it is the interactions within the "structure" of the system that create the "behavior" of that system over time.

An example from another arena often helps people in the pharmaceutical industry understand the concepts around dynamic modeling and risk assessment incorporating uncertainty. The petroleum industry has used these approaches for years to help evaluate the risks and rewards associated with a portfolio of investment opportunities.[6–8] At a 10,000 ft view, the structure inherent in the petroleum industry can be depicted as shown in Figure 35.11.

The rectangles or boxes in Figure 35.11 are called "stocks" and represent places where items (millions of barrels of oil or metrics tons of oil, in this case) can be counted at a single time point. The arrows between the stocks are called "flows" and denote the rate at which items move from one place to another in the system. Note that all these flows are "per time" concepts; *discovery*, *pumping*, and *consumption* are often measured in millions of barrels of oil per year, while *net growth* of the human population would be characterized as people per year. Other icons such as the circles which represent "auxiliary" variables (*total existing oil reserves* and *oil consumption per capita*) and the thin arrows called "links" which represent inter-relationships (human population multiplied by *oil consumption per capita* equals *consumption*) flesh out the set of causal inter-relationships in the system.

The key point about dynamic models is their fundamental operational nature — what we often call the physics of a system. Most people are accustomed to dealing with stocks in the real world, as they are relatively easy to identify and quantify. But operationally, flows are the mechanism through which the magnitude of stocks can change over time. Take another look at Figure 35.11 to see why this is so. We would all like to have a higher quantity of oil in storage tanks — it would provide some peace of mind and might even lower gasoline prices. But the amount of oil in that stock cannot be determined at will. Instead, it is a function of the flows into and out of it. Operationally, the only ways to affect the amount of oil in storage tanks are to:

1. Increase the flow of *pumping* from known oil reserves
2. Decrease the flow of *leakage* which may occur as oil is stored or transferred
3. Decrease the flow of *consumption*

The implication of these various strategies is beyond our purposes here, but it should be evident even to the casual observer that the physics of the system shown in Figure 35.11 shows some long-term problems may be in order. A relatively finite supply of oil (although natural forces are *creating* oil even now, those processes take million of years to occur) coupled with an increasing human population and a propensity for that population to use increasing amounts of oil per capita as developing nations industrialize, point to serious problems in the ability of the oil supply to meet our long-term energy needs.

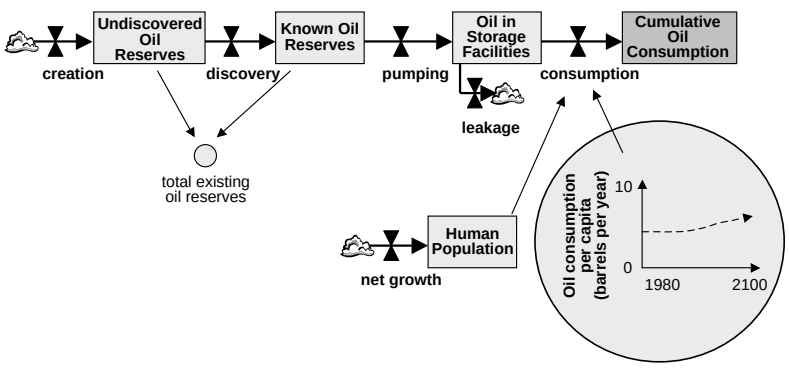


FIGURE 35.11
Dynamic model structure of oil discovery, extraction, and consumption at the industry level.

From the level of an individual petroleum producer, however, the situation is much murkier. Such firms face three key uncertainties when making their strategic plans:

1. The ability to discover new oil fields
2. The volume and costs of oil production in existing and future oil reserves
3. The price at which oil can be sold to consumers

Having a consistent analytic framework that allows these major factors to be considered and the range of possible outcomes to be evaluated is vital to effective long-range planning in the petroleum industry.

In a similar fashion, the stock/flow language of dynamic modeling can be used to describe and ultimately analyze the drug discovery and development process. The basic framework in Figure 35.12 shows the stages through which an NCE must go to reach the marketplace. (The approach in this chapter can also be used for the evaluation of NMEs, but we will reference only NCEs for the sake brevity.) Any of these stages can be expanded to account for more detailed dynamics — the discovery stage, for example, is often divided into subphases specific to an individual company's operating procedures. Figure 35.12 provides a visual, operational backbone of drug development, which can include a number of key dynamics present in pharmaceutical product pipelines. For example, Figure 35.12 depicts failure of NCEs in various stages of development as well as the possibility for retesting of individual compounds.

The flows labeled F1 to F5 represent failures of drugs in various stages of development, including the approval process by the Federal Drug Administration (FDA). Flows F6 and F7 are intended to represent drugs pulled from the marketplace because of the adverse side effects instead of following the normal course of eventual "retirement" through F8. (At least two compounds in the Cox-2 inhibitor drug class were subject to such removal dynamics in 2004 as concerns about their long-term safety were raised.) The clouds represent boundaries

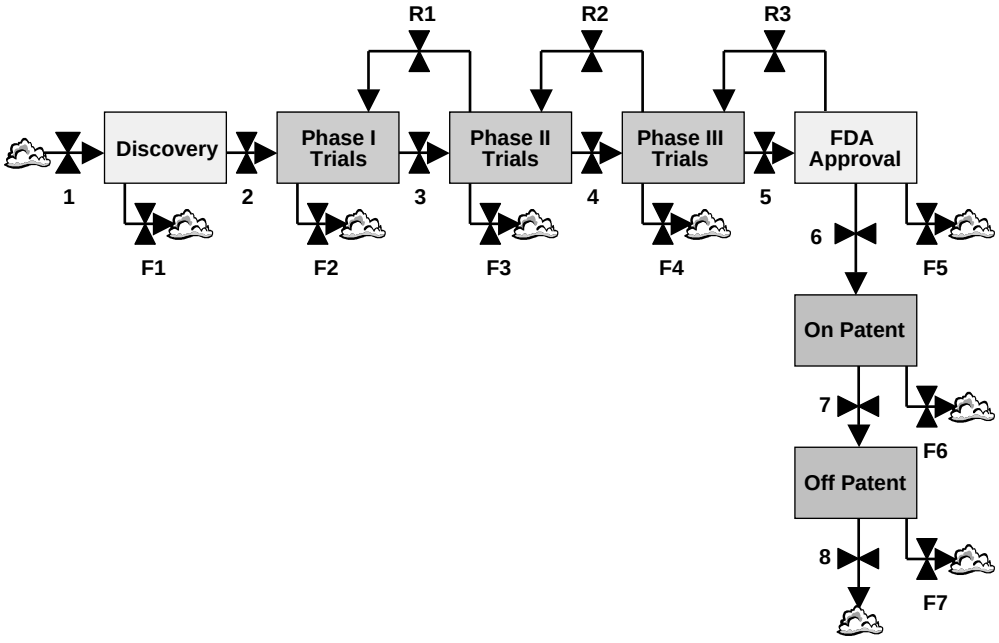


FIGURE 35.12
Drug development framework including compound failures and retesting of phase trials.

for the model, and a model with the stock/flow structure detailed in Figure 35.12 will no longer track any drug compounds entering into a cloud. Because of high compound failure rates, the drug development process is often thought of as a filtering process in which a large number of NCEs in the discovery stage are needed to produce a viable, marketable compound in the on patent phase. The degree of filtering is large, with firms often managing hundreds of NCEs to produce a single revenue-generating market entry. For this reason, pharmaceutical firms often search for and develop multiple NCEs for a particular indication, knowing that most of the compounds will fail at some stage of their development but hopefully at least one will ultimately be successful. Dynamic modeling simulation can help evaluate the risk and uncertainty around such compound filter dynamics in the pharmaceutical pipeline to identify a range of possible outcomes for the pipeline portfolio.

By combining the concept of defined phases through which an NCE must pass with associated failure rates at those stages, dynamic modeling structure provides a consistent visual framework to depict the drug development pipeline. The ability to simulate this structure makes the dynamic modeling approach to pipeline portfolio management even more powerful, as it provides a means to assess the risk associated with a range of possible outcomes.

Some drug failures do not result in complete abandonment of the compound, and in certain cases NCEs may go back through various phases of testing to re-evaluate their potential. The re-evaluation process may identify new indications for the compound on the basis of analysis of results from failed tests. Eli Lilly's drug Strattera, for example, is now used to treat attention deficit/hyperactive disorder (ADHD) but failed clinical trials as a treatment for depression.[9] Problems with dosing, efficacy, and even side effects may dictate a retesting procedure (flows R1–R3), and including such dynamics into the stock/flow framework gives a more complete picture of potential drug developmental pathways.

As Figure 35.12 depicts, "failure" of an NCE is not necessarily a dead end street. The possibility of retesting suggests some human element contributing to phase-test failure — the trial was not designed properly, the metrics proved to be inconclusive, and the dosing regimen was insufficient to show efficacy. Such human errors introduce the inefficiencies of rework into the drug development process, increasing accumulated costs and associated compound development time. Data-driven dynamic model simulation can help identify the impact of the retesting procedures, which represent a strategic opportunity to improve the pipeline simply by reducing introduced errors into the filtering process.

35.6 Populating the Stock/Flow Framework for Drug Discovery and Development

Pharmaceutical companies usually have multiple compounds in various stages of the pipeline, with specific metrics around expected stage completion. The staging of compounds within the drug development pipeline is important for a few reasons:

1. As noted earlier in this chapter, the filtering aspect requires more NCEs in early stages than can be expected to ultimately enter the market.
2. Timing of compounds in various developmental phases affects resource requirements.
3. Eventual market entry and forecasted revenue generation impacts a firm's financial outlook.

For example, a pharmaceutical firm may have information on the status of its drug development process as shown in Table 35.2.

Table 35.2 shows aggregate metrics for time spent per stage, although in reality most firms have expected completion dates specific to individual NCEs. Table 35.3 shows an example of how NCEs in the Phase III process might be categorized in terms of their expected completion dates.

To this point, our modeling examples have not included defined time metrics such as those in Table 35.3, but advances in dynamic modeling simulation technology make this type of data incorporation relatively easy. In addition, a dynamic model can uniquely account for and track individual NCEs as they move through the development process, including compound-specific parameters which govern their eventual behavior. The ability to handle relatively small numbers of items (NCEs, in this case) discretely within the dynamic modeling construct is a powerful way to integrate the overall approach with the details required to analyze drug development portfolios effectively. Some new modeling terminology and concepts must be introduced, however, to extend the explanation of how individual items can be tracked discretely through a dynamic modeling methodology.

35.6.1 Modeling Terminology: Agent-Based Discrete Event Models

Most pharmaceutical firms have pipelines totaling less than 1000 developmental compounds, and often no more than 40. Handling these smaller sample sizes requires an approach, which is capable of attaching individual characteristics to unique, discrete elements. Such models are called "agent-based;" they establish a simulation environment in which individual agents represent actual NCEs in the pipeline, complete with unique characteristics that can be compound-specific. In addition, the small numbers of items tracked dictate that they must be "discrete event;" model inputs must have a range of

TABLE 35.2

Example Data Regarding Drug Development Status and Associated Time Metrics

Stage	Compounds in Stage	Average Time in Stage (Years)
Discovery	20	3
Phase I	12	2
Phase II	10	1.5
Phase III	4	2
FDA approval	2	1.5

TABLE 35.3

Example Product-Specific Completion Date Data for Phase III Compounds

NCE	Expected Phase III Completion Date
Phase III.1	March 2007
Phase III.2	December 2007
Phase III.3	June 2008
Phase III.4	August 2008

probabilities associated with them, the effects of which can be seen through analysis of multiple simulation outcomes.

An agent can be thought of as a stand-alone entity with individual parameters, which govern or track its behavior through the drug development process, as conceptualized in Figure 35.13.

The stock/flow structure in Figure 35.12 can be populated not with generic numbers of compounds and aggregate metrics, but rather with agents representing actual NCEs in the pipeline and their associated individual characteristics. A portion of the overall stock/flow chain, populated with sample agents, is shown in Figure 35.14.

Establishing an agent-based simulation gives the dynamic modeling approach the flexibility and specificity necessary to effectively analyze drug development dynamics consisting of specific unique NCEs. For diagram simplicity, the icons for individual agents will no longer be shown as this example is extended, but they remain the basis for both the modeling approach and all subsequent analysis.

35.6.2 Incorporating Compound Failure Rates

One of the difficulties in pharmaceutical pipeline management lies in quantifying the impact of failure metrics of NCEs in various stages of development. Without some sort of analytical tool to incorporate such risks, the behavior of the overall pipeline is nearly impossible to intuit. Collecting either internal estimates or industry statistics on compound failure rates at different points in the drug development cycle and incorporating them into the dynamic modeling framework can help managers better understand the possible behavior of their current and future drug development

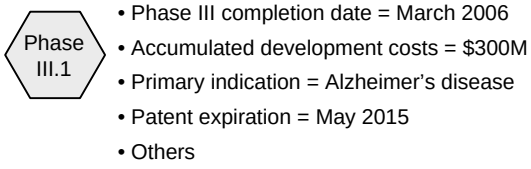


FIGURE 35.13
Conceptual dynamic model simulation agent with associated sample tracking metrics.

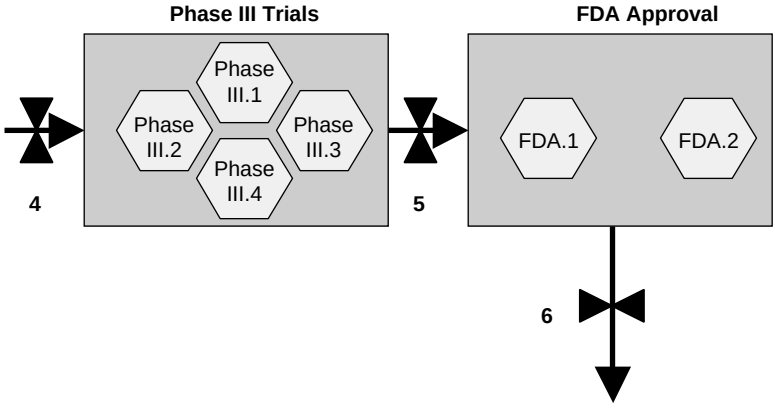


FIGURE 35.14
Portion of the drug development framework populated with sample simulation agents.

process. Data from the Tufts Center for the Study of Drug Development (www.csdd.tufts.edu) are the industry standard for initial estimates regarding compound failure rates, and often provide a base from which a pharmaceutical firm can make an informed judgment regarding the metrics governing the behavior of their unique drug pipeline. For example, a team in charge of pipeline analysis may estimate a set of stage-specific failure probabilities as shown in Table 35.4. (These data are presented in aggregate form, as if these parameters applied to all NCEs in the pipeline. Compound failure rates can be NCE-specific, however, and tied to individual model agents representing a particular compound.)

The estimates in Table 35.4 indicate 60% of drugs in the discovery phase never make it to Phase I (defining the flow of F1). Also, 25% of compounds never make it to the FDA approval phase from Phase III (F4) but an additional 10% have to go backwards in the process to Phase II trials (R2). These data can be incorporated directly into the stock/flow diagram to have a complete picture of the structure and associated parameters for the drug development pipeline, as shown in Figure 35.15. (Note: Some of the parameters governing the NCEs in the pipeline reside at the agent level and are not depicted in this type of aggregate diagram.)

35.6.3 Analysis of Pipeline Dynamics

Even with the fairly simple structure and associated data shown in Figure 35.15, the series of static calculations needed to analyze the possible behaviors of the drug development pipeline over time are overwhelming. However, dynamic modeling software has the capacity to simulate such a model and determine a range of possible outcomes, given the set of input parameters governing the behavior of the system. The most common and insightful simulation result in these types of discrete event models is "Monte Carlo analysis." In a

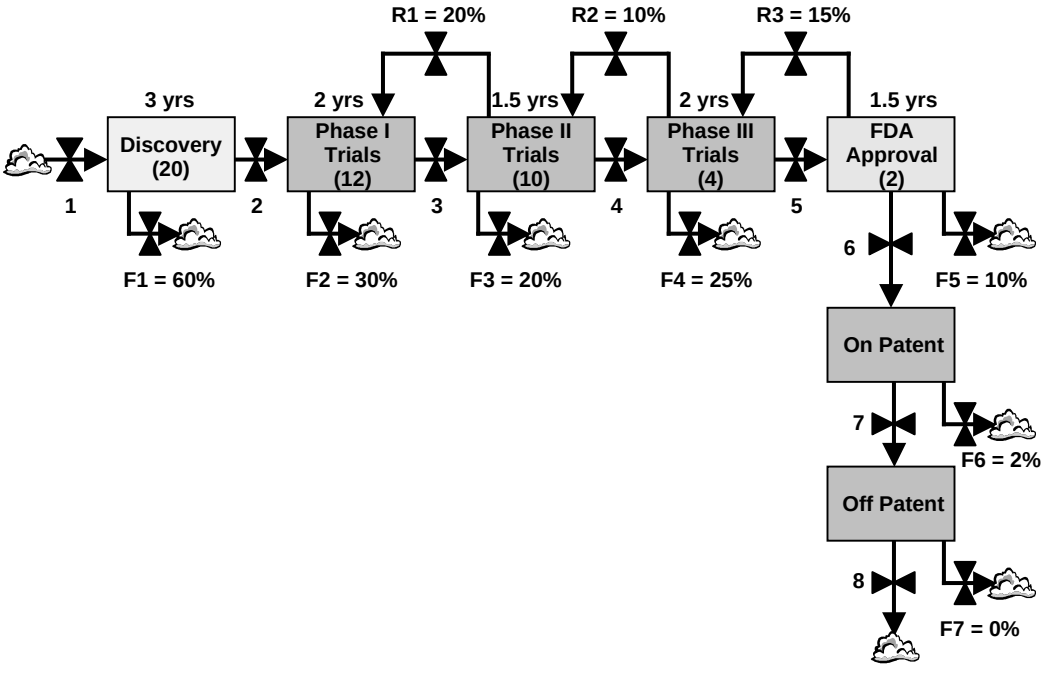


FIGURE 35.15
Drug development framework with associated input data parameters.

Monte Carlo experiment, a large number of simulations are performed independently and the corresponding results are grouped together for analysis. In effect, each simulation outcome is unique as it represents the results of a specific set of input assumptions. But these input assumptions are drawn from a distribution and attached to individual agents (representing actual NCEs) for each simulation, as in the case of compound failure rates. For example, in 300 of 1000 simulations would a particular NCE fail the Phase I trial stage (30% — see Table 35.4). The individual agents of the dynamic model are parameterized, as the beginning of each run of a multiple simulation Monte Carlo analysis, producing a range of possible outcomes. In this manner, Monte Carlo analysis provides an analysis of the overall variability of model results on the basis of a stock/flow structure and a defined set of discrete event input assumptions attached to individual agents.

Monte Carlo analysis can also incorporate probability distributions for input parameters. As shown in Table 35.2 and Table 35.4, first-pass parameters for dynamic pipeline portfolio models are often assumed to be a single value, such as the average time in the discovery process being 3 years. However, more advanced analysis of new product development dynamics often includes a range around these mean values, introducing yet another layer of uncertainty and associated realism into the resulting dynamic simulations. For example, the parameter governing the time spent in the discovery phase can be established as a probability distribution, with a range of possibilities around a mean of 3 years, as depicted in Figure 35.16.

TABLE 35.4

Example Data for Stage-Specific Compound Failure Rates

Stage	Variable Name	Failure Rate (%)
Discovery	F1	60
Phase I	F2	30
Phase II	F3	20
Phase II	R1	20
Phase III	F4	25
Phase III	R2	10
FDA approval	F5	10
FDA approval	R3	15
On patent	F6	2
Off patent	F7	0

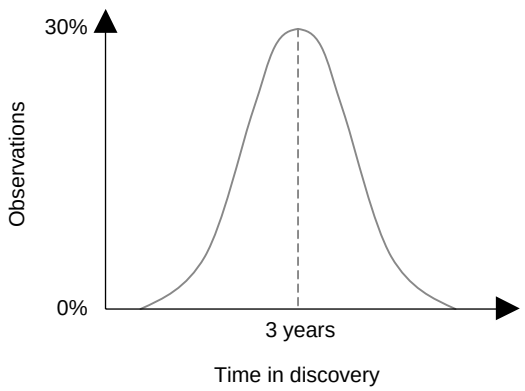


FIGURE 35.16

Establishing input parameter probability distribution (e.g., time in discovery).

For each simulation and for each agent representing an NCE, the dynamic model will choose a value from the distribution to govern the time that particular NCE spends in the discovery phase. Including this dimension of variability can often affect the overall distribution of potential outcomes as determined by Monte Carlo analysis, and creates a simulation environment that more closely mimics the real world of drug development with uncertain governing parameters.

The determination of appropriate probability distributions and their ultimate impact on overall portfolio dynamics is beyond the scope of this work, but dynamic modeling software is flexible enough to include a wide range of choices in this regard. Inclusion of uncertainty around input parameters is easy to incorporate into dynamic modeling simulations, and often adds a degree of robustness to overall model methodology and corresponding outputs.

To continue with this example, assume that the flow into the discovery stage (flow *1*) is ten NCEs per year. In other words, a pharmaceutical company expects its scientists to identify ten new compounds, which will enter the discovery process each year. Given this input to the system and the overall times associated with stage completion and associated failure rates, what is the range of outcomes of the firm's new product pipeline over time? This example will assume that no drugs are currently in the on patent or off patent stages, although including such on-market compounds is an easy extension.

Figure 35.17 shows the most likely outcomes for the firm's drug discovery and development process on the basis of a Monte Carlo analysis of the dynamic model established in this chapter.

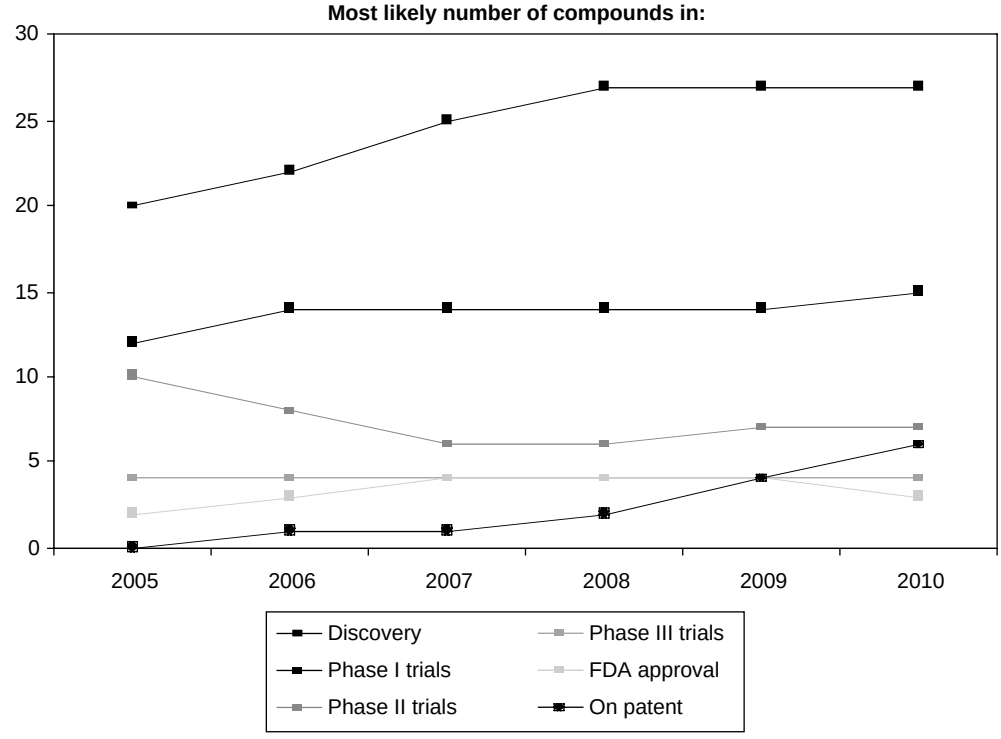


FIGURE 35.17
Example trajectory of most likely number of compounds in various development stages.

Charts such as the one in Figure 35.17 are useful in analyzing the likely evolution of a pharmaceutical firm's product portfolio over time. For example, it shows that the number of NCEs in Phase III trials is expected to stay constant from 2005 through 2010. However, the number of drugs in discovery will likely increase dramatically over that same interval. These patterns give aggregate evidence as to likely evolution of the new product pipeline, and often suggest associated resource requirements and medium-range financial planning issues.

As noted earlier in this chapter, however, "most likely" does not imply certainty, and a more illuminating result from a Monte Carlo analysis is the range surrounding an expected result. For example, the projection of the number of compounds on patent can be shown with a distribution around expected outcomes as shown in Figure 35.18.

Although the results for the on patent stage are shown in Figure 35.18, the dynamic model can capture the resulting data from any of the defined phases in the stock/flow diagram. The results of this type of Monte Carlo simulation can be analyzed to produce a series of useful metrics, such as:

1. Confidence intervals: What percentage of results fall above or below a certain threshold?
2. Standard deviation: How wide is the distribution of results around the expected or most common outcome?
3. Causes of variability: Which parameters, such as average time spent in each stage or compound failure probabilities at various phases, contribute the most to the distribution of overall outcomes?
4. Sequencing of adverse outcomes: What combinations of factors contribute to the outcomes which differ the most from the expected result?

A careful examination of these types of analyses leverages the dynamic modeling simulation methodology with a Monte Carlo approach, and can help portfolio managers better understand how their new product development process might evolve over time. This more complete understanding can then be translated into better strategic decisions regarding portfolio management and the strategic decisions surrounding a firm's developmental pipeline.

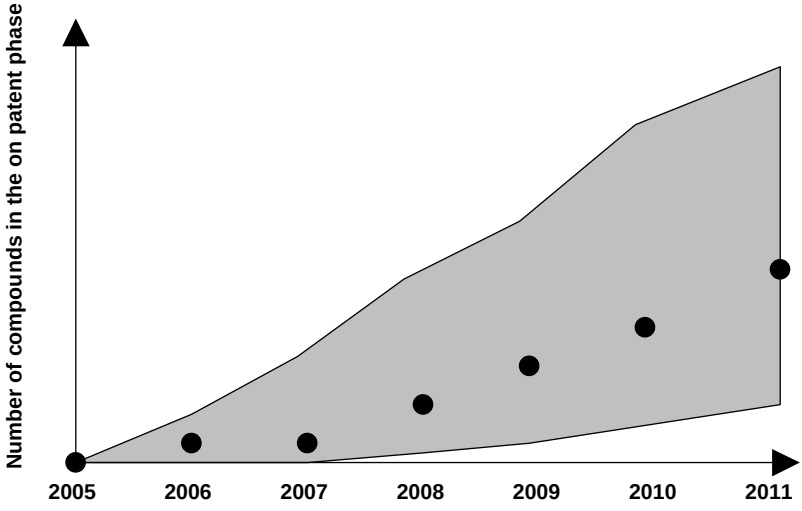


FIGURE 35.18
Example risk trajectory for number of compounds on patent.

35.7 Dynamic Modeling vs. Alternative Pipeline Portfolio Analysis Tools

Some readers may have seen pharmaceutical portfolio analysis designed and implemented in a spreadsheet or other computational applications, including the ability to compute Monte Carlo results. While these approaches can indeed be of value, our clients have seen a number of additional insights into pipeline dynamics, which *only* a dynamic modeling simulation can provide.

35.7.1 Advantages of Dynamic Modeling in Drug Discovery and Development Analysis

1. Dynamic models are more flexible in their establishment of structure/metrics, and are not restricted to the financial metaphor underlying many other approaches.
2. Stock/flow methodology provides a visual framework to conceptualize the new product development process.
3. Operational diagram can be populated with simulation agents representing individual compounds in various developmental stages.
4. Agents are easily aggregated, extended, and duplicated, making dynamic models easy to appropriately scale and subsequently analyze.
5. Compound-specific attributes can be assigned to agents, which may then be inherited by LEs generated from parent compounds.
6. Interaction between NCEs is easily captured, as in cases where success/failure of one compound influences the expected success/failure of another.
7. Dynamic model structure can be simulated to determine the expected outcomes from a set of associated input assumptions, both in terms of existing pipeline and future portfolio scenarios.
8. Condition of the pipeline can be observed at any time point by pausing the simulation and observing the status of the system.
9. Dynamic models explicitly categorize the causes for model results instead of simply calculating combinations of statistical distributions. Analysis of simulation results can categorize not only *what* happened but also *why* it happened.
10. Stock/flow structure combine uncertainty with cause-and-effect relationships which can be quantified through discrete event simulation.
11. Dynamic simulations can categorize system variability, including the effect of strategic decisions on changes in metrics over time.
12. Supply/production issues can be combined with pipeline evaluation to provide an integrated tool for scenario planning.

Many of these advantages will be demonstrated throughout this chapter as the basic dynamic modeling framework for analysis of drug discovery and development is extended.

35.8 Extensions of the Basic Dynamic Modeling Drug Development Structure

The basic stock/flow structure and the associated simulation model of pharmaceutical pipeline management provide a useful framework to analyze basic drug development

processes. Readers familiar with these issues, however, know that the scope of actual pipeline portfolio management is much more complex. Fortunately, the operational simulation framework is flexible and well suited for some key additions. Specifically, five separate extensions have proven themselves valuable to our clients over the years. These extensions will be incorporated in isolation, or integrated to develop a comprehensive dynamic model of the entire pipeline portfolio management process.

35.8.1 Common Extensions of the Basic Stock/Flow Structure of Drug Development

1. Taking a compound to OTC market
2. In-licensing compounds from other firms
3. Out-licensing compounds to other firms
4. Including LEs from parent compounds
5. Computing resource requirements

The first four of these common extensions are depicted in Figure 35.19.

35.8.2 Taking a Compound to the Over-the-Counter Market

Firms with compounds either currently or about to go off patent are increasingly turning to the OTC market as a potential strategic option for compound lifecycle management. Advil is perhaps the most successful of these OTC transitions, but in recent years brands such as Claritin, Nicorette, and Benadryl have made similar conversions. The details of evaluating, managing, and implementing are myriad, but a simple extension of the basic stock/flow framework provides a means to test and analyze the revenue effects of this strategy. Figure 35.19 shows how the dynamics of the move to OTC can be captured in dynamic modeling methodology of drug development.

The auxiliary variable *OTC revenues per compound* is of course not a simple constant metric, and a complementary analysis is often done to determine the scope and magnitude of the factors which affect it. Some of these influencers might be:

1. Size of the associated OTC market
2. Number and aggressiveness of existing competitors
3. Effectiveness of the compound in the indication
4. Marketing spending in support of the compound while on the OTC market

A comprehensive analysis of these factors is beyond the scope of this work, but these influencers are often incorporated into the dynamic model to more fully address the revenue implications of a move to the OTC market for a given compound.

35.8.3 In-Licensing Compounds

Pharmaceutical firms often choose to in-license compounds to beef up their developmental pipeline or round out their existing product portfolio. Structurally, in-licensing can be represented with a simple extension of the basic stock/flow framework, as shown in Figure 35.19.

In-licensing opportunities can enter the drug development process at the level of phase trials and the on patent stage. In either case, the simulation model can assign a cost necessary to obtain in-licensed compounds (*cost per in-licensed compound*), which can be compared

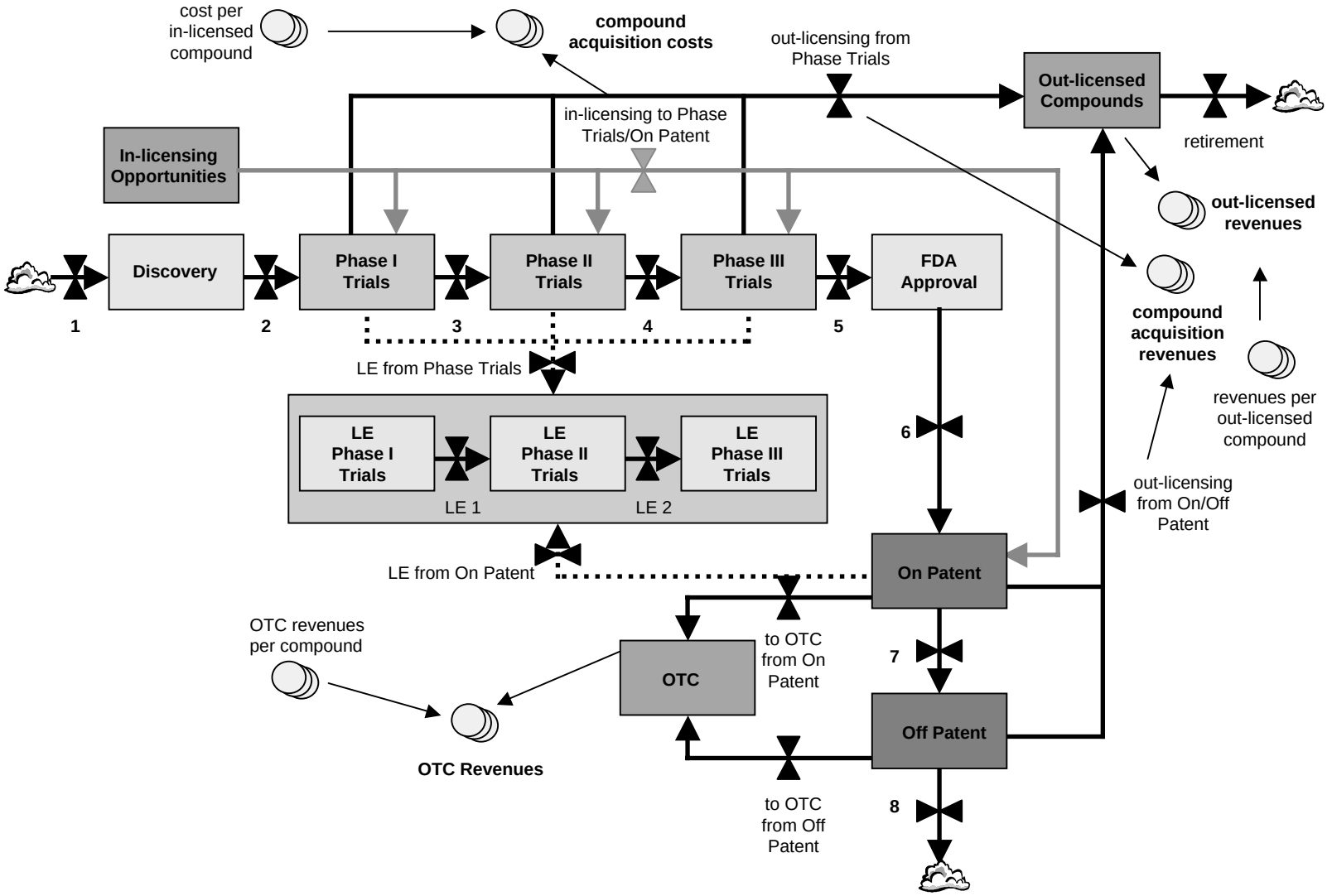


FIGURE 35.19
Expanded drug development framework with strategic options for pipeline portfolio management.

to their expected revenue streams to determine the financial return of these types of in-licensing strategies. Once entering the stock/flow model, in-licensed compounds are assigned dynamic metrics, which govern their subsequent behavior. These parameters can be the same as those for internally generated NCEs, or in-licensed compounds may be given differential probabilities and tracked separately as they advance through the development pipeline. Some of our clients, for example, assign higher failure probabilities for in-licensed compounds, reflecting a higher inherent level of uncertainty regarding their chances of success. Monte Carlo analysis, as described previously in this chapter, is then performed to determine the range of possible outcomes for this set of in-licensed opportunities.

The timing of in-licensing is an important aspect in evaluating potential external opportunities. In a world of diminishing pipelines, the in-licensing of late-stage compounds has become an expensive game to play. As pharmaceutical firms are forced to evaluate in-licensing opportunities further and further back in the development pipeline, the need for careful analytics expands. One nice feature of the dynamic modeling approach is the ability to bring external NCEs into the simulated pipeline structure at any desired stage of development, and subsequently subject those agents to specific parameters governing their behavior.

35.8.4 Out-Licensing Compounds

Out-licensing opportunities represent another strategic option for firms to decrease the number of compounds in their pipeline and generate immediate cash by "exporting" compounds to other firms. The possible revenue generation from and the strategic nuances to out-licensing are varied, and can be analyzed with the help of an extension to the basic stock/flow framework as shown in Figure 35.19.

Typically, revenue streams from such divestitures come in two forms:

1. An acquisition fee paid by the acquiring company for the right to develop the acquired compound (*compound acquisition revenues*).
2. Some sort of revenue-sharing arrangement if and when the compound reaches the market (*out-licensed revenues*.) The variable *revenues per out-licensed compound* is commonly a time-based scenario of revenue generation over the life of the compound before it eventually exits through the *retirement* flow. A proper analysis of the *revenues per out-licensed compound* flow should incorporate the uncertainty of market entry, however, if the out-licensed NCE is one which has not yet passed the FDA approval process.

Pharmaceutical firms are understandably hesitant to out-license their compounds, often because of lack of clarity regarding the strategic and financial implications of such actions. Fear of creating competition in an important indication, unwillingness to depart with viable compounds, and uncertainty around the ultimate impact of out-licensing decisions certainly make the case for caution on the part of portfolio managers. Including out-licensing possibilities into the dynamic simulation model, however, provides the means to analyze and quantify the potential pros and cons of such options in a consistent and clear manner. The uncertainty surrounding out-licensing will never disappear, but the analytical framework of a dynamic pipeline portfolio model can be used to comprehensively evaluate this strategic option.

35.8.5 Line Extensions from Parent Compounds

During the development process or even while in the on patent stage, drug compounds sometimes develop what are called LEs: new indications, methods of administration, or

even dosing regimens which create new business opportunities for existing compounds. Line extensions generally have to go through the same testing phases as NCEs, although they may not necessarily have to begin at Phase I. Because LEs are "offshoots" of parent compounds, they do not replace the NCEs from which they are spawned. Instead, LEs should be treated as newly generated compounds, which retain some of the properties of their parent compound (patent expiration, accumulated costs, failure of the parent compound triggering failure of the LE, etc.) Incorporating these dynamics is fairly straightforward from a structural standpoint, as shown in Figure 35.19.

The dotted lines on the flows of *LE from On Patent* and *LE from Phase Trials* in Figure 35.19 denote that existing, parent compounds are not actually physically transported back to LE phase trials when an LE is created. Instead, a new agent is created in the model that retains certain properties from the parent agent, and this new agent is then sent along the path of LE phase trials. The subsequent behavior of the LE will be governed by how this new simulation agent is parameterized. Simulation of the LE dynamics gives a sense for the timing and implication of their inclusion in the pipeline process in a way that static approaches have difficulty capturing. In particular a pharmaceutical world where discovery of viable new compounds appears to be slowing, accurately portraying the dynamics associated with LEs will be increasingly important.

To fully incorporate the dynamics of LEs, a number of assumptions must be agreed upon and implemented in the dynamic modeling structure. These LE parameters often include, but are not limited to:

1. How often does a parent compound generate an LE?
2. What testing phase does an LE enter?
3. What are the residence times and failure rates for an LE in phase trials?
4. Can an LE generate extensions from itself?
5. Does an LE fail if the parent compound fails any of its testing phases or the FDA approval stage?

The inclusion of LE structure and the ability to quantify the parameters, which govern its behavior provides a more realistic evaluation of drug development pipeline dynamics over time. The agent-based nature of dynamic modeling simulations makes this extension fairly easy to implement, and allows for LEs to be analyzed as a distinct subset of the entire portfolio management process. The flexibility of this methodology allows for a very operational representation of the LE possibilities and seamlessly quantifies their inclusion in the overall drug pipeline.

35.8.6 Computing Resource Requirements

The uncertainty contained in the pharmaceutical development process makes the associated resource planning task a daunting one. However, dynamic modeling simulations can shed some light on the resource requirement planning process by quantifying a range of possible outcomes from a firm's new product pipeline. These outcomes, usually measured in terms of number of compounds at various developmental stages over time, can be easily translated into the number of resources necessary over a certain time horizon. For example, the sample outputs shown in Figure 35.18 can be converted into resource requirements on the basis of expected number of personnel needed per compound at any given development state.

Such resource requirement projections often identify important strategic gaps, which may require immediate attention. For example, large increases in estimated full-time

equivalents (FTEs) implied by likely evolution of a particular new product development process may not be possible, much less practical, given the current labor market. This analysis is often extended to include dynamics of up-to-speed time and productivity metrics associated with such rapid hiring. One possible strategy to manage this expected FTE gap is to ramp up hiring in the earlier years in anticipation of increased resource needs in the longer term. Alternatively, an out-licensing strategy can be tested to reduce internal resource demands. By providing an integrated means to determine a firm's resource needs over time, the dynamic portfolio simulation allows managers to systematically test and evaluate various strategic solutions and make informed decisions in a timely manner.

35.8.7 Strategy Tests Using the Dynamic Model Simulation

Dynamic models such as those described in this chapter can be used for a wide variety of scenario testing, what-if analysis, risk assessment, portfolio optimization, and strategy development. Our clients find these analyses both insightful and actionable, combining the power of well-established principles from the world of financial services with the power and flexibility of the dynamic modeling approach to portfolio management. The result is a framework that addresses the subtleties of pharmaceutical discovery and development — nuances that must be recognized and accounted for in effective risk assessment and resulting strategies related to portfolio planning.

35.9 Part 2 Recap

1. Dynamic modeling methodology provides a unique way to address portfolio management problems in the pharmaceutical industry.
2. The framework establishes a stock/flow structure representing stages that an NCE must pass so as to reach the marketplace.
3. Associated failure metrics can be assigned to NCEs, either on an aggregate or individual compound level.
4. Agent-based models create individual entities in the dynamic modeling framework to represent unique chemical or biologic compounds.
5. Basic stock/flow structure related to drug development can be extended or modified to address various strategic options for NCE management.
6. Dynamic modeling simulations incorporate a variety of analytic techniques to identify uncertainty and assess resulting risk in the drug discovery and development processes.

References

1. Landers, P., Cost of developing a new drug increases to about $1.7 billion, *Wall Street J.*, December 8, 2003.
2. Richmond, B. and Peterson, S., *An Introduction to Systems Thinking*, from the users manuals for the ithink® software, 1996. Available from iseesystems, Inc., Lebanon, NH.
3. Lilien, G., Kotler, P., and Moorthy, K.S., *Marketing Models*, Prentice-Hill, Englewood Cliffs, NJ, 1992.

4. Pharmaceutical Industry Profile, Pharmaceutical Research and Manufacturers of America (PhRMA), 2002.
5. On Treacherous Ground, 50 of the world's biggest pharmaceutical companies are challenged to perform in the most difficult environment ever, *Med. Ad News*, September 1, 2002.
6. Haspeslagh, P., Portfolio planning: Uses and limits, *Harvard Business Review*, January–February, 1982, p. 58.
7. Hightower, M.L. and David, A., Portfolio modeling: a technique for sophisticated oil and gas investors, Society of Petroleum Engineers, SPE 22016, Presented at the SPE Hydrocarbon Economics and Evaluation Symposium, Dallas, Texas, 11–12 April 1991.
8. Ball, B.C. and Savage, S.L., Portfolio thinking: beyond optimization, *Petroleum Eng. Int.*, 54–56, 1999.
9. Burton, T., By learning from failures, Lilly keeps drug pipeline full, *Wall Street J.*, April 21, 2004.