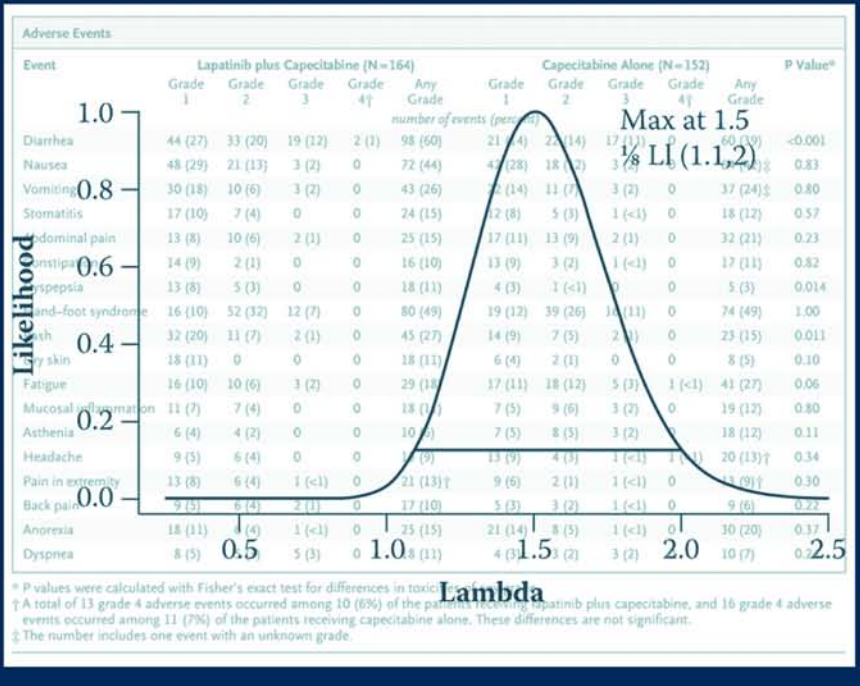
Chapman & Hall/CRC Biostatistics Series

Data and Safety Monitoring Committees in Clinical Trials



Jay Herson

CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK

Data and Safety Monitoring Committees in Clinical Trials

Chapman & Hall/CRC Biostatistics Series

Editor-in-Chief

Shein-Chung Chow, Ph.D.
Professor
Department of Biostatistics and Bioinformatics
Duke University School of Medicine
Durham, North Carolina, U.S.A.

Series Editors

Byron Jones
Senior Director
Statistical Research and Consulting Centre
(IPC 193)
Pfizer Global Resea ch and Development
Sandwich, Kent, U. K.

Jen-pei Liu
Professor
Division of Biometry
Department of Agronomy
National Taiwan University
Taipei, Taiwan

Karl E. Peace
Georgia Cancer Coalition
Distinguished Cancer Scholar
Senior Research Scientist and
Professor of Biostatistics
Jiann-Ping Hsu College of Public Health
Georgia Southern University
Statesboro, Georgia

Bruce W. Turnbull
Professor
School of Operations Research
and Industrial Engineering
Cornell University
Ithaca, New York

Chapman & Hall/CRC Biostatistics Series

Published Titles

1. *Design and Analysis of Animal Studies in Pharmaceutical Development,* Shein-Chung Chow and Jen-pei Liu
2. *Basic Statistics and Pharmaceutical Statistical Applications,* James E. De Muth
3. *Design and Analysis of Bioavailability and Bioequivalence Studies, Second Edition, Revised and Expanded,* Shein-Chung Chow and Jen-pei Liu
4. *Meta-Analysis in Medicine and Health Policy,* Dalene K. Stangl and Donald A. Berry
5. *Generalized Linear Models: A Bayesian Perspective,* Dipak K. Dey, Sujit K. Ghosh, and Bani K. Mallick
6. *Difference Equations with Public Health Applications,* Lemuel A. Moyé and Asha Seth Kapadia
7. *Medical Biostatistics,* Abhaya Indrayan and Sanjeev B. Sarmukaddam
8. *Statistical Methods for Clinical Trials,* Mark X. Norleans
9. *Causal Analysis in Biomedicine and Epidemiology: Based on Minimal Sufficient Causation* Mikel Aickin
10. *Statistics in Drug Research: Methodologies and Recent Developments,* Shein-Chung Chow and Jun Shao
11. *Sample Size Calculations in Clinical Research,* Shein-Chung Chow, Jun Shao, and Hansheng Wang
12. *Applied Statistical Design for the Researcher,* Daryl S. Paulson
13. *Advances in Clinical Trial Biostatistics,* Nancy L. Geller
14. *Statistics in the Pharmaceutical Industry, Third Edition,* Ralph Buncher and Jia-Yeong Tsay
15. *DNA Microarrays and Related Genomics Techniques: Design, Analysis, and Interpretation of Experiments,* David B. Allsion, Grier P. Page, T. Mark Beasley, and Jode W. Edwards
16. *Basic Statistics and Pharmaceutical Statistical Applications, Second Edition,* James E. De Muth
17. *Adaptive Design Methods in Clinical Trials,* Shein-Chung Chow and Mark Chang
18. *Handbook of Regression and Modeling: Applications for the Clinical and Pharmaceutical Industries,* Daryl S. Paulson
19. *Statistical Design and Analysis of Stability Studies,* Shein-Chung Chow
20. *Sample Size Calculations in Clinical Research, Second Edition,* Shein-Chung Chow, Jun Shao, and Hansheng Wang
21. *Elementary Bayesian Biostatistics,* Lemuel A. Moyé
22. *Adaptive Design Theory and Implementation Using SAS and R,* Mark Chang
23. *Computational Pharmacokinetics,* Anders Källén
24. *Computational Methods in Biomedical Research,* Ravindra Khattree and Dayanand N. Naik
25. *Medical Biostatistics, Second Edition,* A. Indrayan
26. *DNA Methylation Microarrays: Experimental Design and Statistical Analysis,* Sun-Chong Wang and Arturas Petronis
27. *Design and Analysis of Bioavailability and Bioequivalence Studies, Third Edition,* Shein-Chung Chow and Jen-pei Liu
28. *Translational Medicine: Strategies and Statistical Methods,* Dennis Cosmatos and Shein-Chung Chow
29. *Bayesian Methods for Measures of Agreement,* Lyle D. Broemeling
30. *Data and Safety Monitoring Committees in Clinical Trials,* Jay Herson

Chapman & Hall/CRC Biostatistics Series

Data and Safety Monitoring Committees in Clinical Trials

Jay Herson
The John Hopkins University
Baltimore, Maryland, U.S.A.

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

A CHAPMAN & HALL BOOK

Cover art from Geyer, C.E., Forster, J., Lindquist, D. et al. (2006), *New England Journal of Medicine,* **355**, 2733–2743. With permission. Copyright © 2006, Massachusetts Medical Society. All rights reserved.

Chapman & Hall/CRC
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

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-13: 978-1-4200-7037-8 (Hardcover)

Library of Congress Cataloging-in-Publication Data

Herson, Jay.
Data and safety monitoring committees in clinical trials / Jay Herson.
p. ; cm. -- (Chapman & Hall/CRC biostatistics series ; 30)
Includes bibliographical references and index.
ISBN 978-1-4200-7037-8 (hardcover : alk. paper)
1. Drugs--Testing--Evaluation. 2. Drugs--Testing--Statistical methods--Evaluation. I. Title. II. Series: Chapman & Hall/CRC biostatistics series ; 30.
[DNLM: 1. Clinical Trials Data Monitoring Committees--organization & administration. 2. Clinical Trials as Topic--standards. 3. Data Collection. 4. Drug Industry--standards. 5. Research Design. 6. Safety Management. QV 771 H572d 2009]

RM301.27.H47 2009
615'.19--dc22 2009003127

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Dedication

To the patients who volunteer for clinical trials—past, present and future.

Contents

Preface

This is a book about *best practices* in safety monitoring through data monitoring committees (DMCs) in pharmaceutical industry clinical trials. It will be useful for those who have served on DMCs, those interested in what was done well and what could have been done better, and those contemplating serving on their first committee.

I can still remember that winter morning in San Antonio twenty years ago when Frank Rockhold asked me at a biostatistical meeting we were attending if I could organize a DMC for a gastrointestinal drug his company was developing. Frank, his talented colleagues, and I worked out a plan and the first DMC in pharmaceutical industry trials was born. At the time I headed a contract research organization (CRO) in Houston known as Applied Logic Associates (ALA). Since then that company has provided statistical support to more than 50 DMCs and I have served as statistical member on DMCs for about 30 additional trials. The art and science of safety monitoring through DMCs have reached adolescence and it is time to review, and perhaps debate, best practices. There can be no better time than now, when regulatory agencies worldwide are facing considerable challenges in drug safety for both premarket and postmarketed drugs.

In the world of drug development, clinical issues and statistical issues cannot be separated. All issues are scientific. All use *applied logic*. This is the approach to this book. It is written in the style of my "Herson's Handout" column that appeared in the ALA newsletter *Under the Curve*, 1991–2004. This was a style appropriate for all drug development professionals regardless of degrees held. The book assumes that the reader has a basic knowledge of clinical trials, clinical operations, and good clinical practices.

Chapter 1 is introductory. It points out the differences between clinical trials sponsored by the federal government and those sponsored by pharmaceutical firms. These differences explain why DMCs operate a little differently for private sector–sponsored trials. We learn that pharmaceutical companies themselves differ by their size, and the terms *Big Pharma*, *Middle Pharma*, and *Infant Pharma* are defined. Another important definition is that of stewardship. If there had to be one word to define the role of DMCs, it would be *stewardship*. We also learn the limitations for uncovering safety issues in a single premarket clinical program and provide a rule of thumb for assessing the sensitivity of a clinical program to uncovering adverse events.

Chapter 2 details the organization of safety monitoring describing the interactions of the sponsor, the DMC, the Data Analysis Center (DAC), the Institutional Review Board (IRB), and the regulatory agencies. Tables in the chapter offer checklists of desirable characteristics of a sponsor representative and a DAC organization, and questions to ask oneself and the sponsor before agreeing to join a DMC.

Chapter 3 explains the nature of DMC meetings. Of special importance is the orientation meeting. In this section the items usually included in the DMC charter are listed. From this the extent of DMC responsibilities for the trial, reporting procedures, serious adverse event data flow, masking (blinding) policy and many other important agreements that must be made between the sponsor, DAC, and DMC at the outset of the trial are detailed in text and tables. Another table presents a sample agenda format for DMC meetings that has proven more useful than a mere list of topics to be discussed.

Chapter 4 is an introduction to clinical issues. Here we see how the safety data reviewed by DMCs arise. We learn the important distinction between adverse events, serious adverse events, and severe adverse events. The current state of the art of adverse event coding is described. The impact of multinational trials and the cultural, political, and medical practice issues relevant to DMC operations are described.

In Chapter 5 we investigate statistical methods useful for DMCs. It is emphasized that statistical significance of a treatment difference for a safety parameter is neither a necessary nor a sufficient reason to terminate a trial. We see some useful graphical and tabular data displays and review statistical methods for testing hypotheses and creating confidence intervals for various measures of treatment differences in safety. The methods are illustrated with data from an actual clinical trial. Every DMC faces problems of multiplicity, and this concept is explained and use of the false discovery rate is presented as a means of controlling multiplicity. The chapter includes an introduction to likelihood methods for assessing evidence. Much of the work in this area has been done by my graduate school advisor, Professor Richard Royall. I have found the methods useful for DMC work. I am excited to be able to present the methods here. The chapter closes with a brief description of the role of Bayesian methods for safety analysis. A table is presented summarizing all methods discussed in the chapter along with their advantages and disadvantages for DMC use.

Chapter 6 continues in the inference vein with a description of the biases and pitfalls in analyzing safety data. Sources of bias for arising from unmasking, incomplete follow-up, spontaneous versus solicited adverse event data collection, early termination due to efficacy, and the problems introduced by slicing and dicing the adverse event descriptions into many subgroups within a body system (granularity) so that statistically significant events can be detected or missed. Finally the concept of competing risks in adverse event incidence is presented, particularly in the case where there is differential follow-up between treatment groups due to a treatment effect in a primary efficacy endpoint.

We now arrive at Chapter 7, where we apply our knowledge from prior chapters to data monitoring committee decisions. We review the types of decisions DMCs can make as well as the environment in which they are made. We see the steps that can be taken when a safety issue arises and the potential pitfalls in incorporating data from past clinical trials into the decision-making process.

Chapter 8 might also be called an epilogue. It deals with emerging issues in drug development that affect DMC operations. The issues are divided into those that arise from advances in technology and those that arise through the maturation of the DMC process. In the former we take up adaptive designs. Of particular importance is the situation when an adaptive change is taking place on the basis of efficacy but, due to

safety, the DMC feels that this change will not be in the interest of patient safety. Other technology changes are the advent of real time SAE reporting and the potential of certain adverse events to be biomarkers for efficacy. In the area of maturation of DMC processes, we observe that the training of DMC members, CROs, DACs, and even sponsor representatives to the DMC paradigm is very important. We consider how we can create the supply of qualified individuals to meet the demands. Sponsors have encountered problems with cost control as DMCs ask for more data than was originally planned. Suggestions are given for dealing with this situation. What happens when pharmaceutical companies merge or out-license the product the DMC is deliberating? How do we ensure that independent review of patient safety will continue? New medical journal policies requiring DMCs and independent statistical review will also affect DMC operations and are covered here as well.

At the end of each chapter the reader will find a Q & A section called DMCounselor. The questions provide a behind-the-scenes glimpse of DMC meetings, interactions with sponsors, multinational issues, personality conflicts, and especially the problems that Infant Pharma faces in providing independent review in the same manner as Big Pharma. All cases raised in this section are real. They were either my own personal experiences or those of others. All details of the cases have been changed for confidentiality but the conundrums remain. Some readers may not agree with the solutions that I provide in my answer, but at least they can see issues they may not have thought of previously.

A glossary is presented giving definitions of most of the technical terms used in the text, and the appendix contains a table of adverse events reported for selected marketed drugs in placebo-controlled trials. It is referred to at various points in the text to illustrate safety concepts.

I explain in the text that I use the term *sponsor* where others might use *company.* I use the term *patient* as the clinical unit in our trials. I realize that in some indications *subject* would be more appropriate but I chose patient for consistency. I use the term *drug* as a synonym for *intervention*. The latter term would include biologics and medical devices. I felt that most readers would be used to this terminology. Indeed the Food and Drug Administration and the Drug Information Association deal with interventions broader than just drugs.

About a year after completing my first statistics course as an undergraduate, I noticed that the professor had published his own textbook on the subject. When I asked him how long it took him to write the book, he replied, "Well, it is hard to say. I have been teaching for twenty years." I now understand what he meant. Although writing this book may have taken one year, it represents twenty years of experience. This also means that it will be impossible for me to thank all those people from whom I have learned. I certainly must thank Frank Rockhold for introducing me to the DMC concept. I thank the many employees of ALA, too numerous to name here, for their insight into the details, such as MedDRA and related software, their development of efficient methodology, and the realization that protecting patient safety was just as important as hitting a home run in efficacy. During my years as president of ALA I became intimately aware of the issues facing Infant Pharma and am grateful for the opportunity to present them here. There have been many DMCs. If I had to single out one DMC experience, it would be the one that lasted longest. My hat goes off

to ophthalmologists Alan Bird, Don D'Amico, and Ron Klein; DAC biostatistician Emmanuel Quinaux; and sponsor representative Harvey Masonson for all they taught me while we had the privilege of working on the first vascular endothelial growth factor used in ophthalmology. It was a very educational six years.

I am so grateful for the Biostatistics Department at the Johns Hopkins Bloomberg School of Public Health in Baltimore for welcoming me into the department after my semiretirement and offering me the physical and intellectual environment conducive to writing a book. Among the Johns Hopkins family, Scott Zeger brought me into the department in 2004 and Richard Royall not only introduced me to likelihood methods when I was his student but also set the example for me by writing such a fine book on the subject. Jeff Blume and Elizabeth Garrett-Mayer also shared a lot of ideas about likelihood and helped with the graphs.

I thank biostatisticians Dennis Dixon, Janet Wittes, and Marc Buyse not only for our long friendship but also for teaching me the differences between government and private industry trials. I thank Janet also for her official responsibility of reviewing this manuscript and making so many important suggestions and corrections.

This book would never have come about if David Grubbs of the Taylor & Francis Group had not approached me about writing a book and offered encouragement along the way. Taylor & Francis could not have provided me with a better project coordinator than Marsha Pronin.

Last, but certainly not least, I must thank my wife, Linda, for enduring my writing obsession and providing so much encouragement for what is the equivalent of a second doctoral dissertation some thirty-seven years after the first. Yes, Linda, I'll clean up the room now.

I hope that the many drug development professionals who read this book will find it useful in starting a dialogue on best practices beyond the point of just writing a charter and scheduling meetings. I hope that those talented readers who have not yet served on a DMC will consider doing so. It is an important and rewarding experience.

Jay Herson, Ph.D.

Chapter 1

Introduction

Bullets to Remember

- Data monitoring committees are responsible for the *stewardship* of a clinical trial.
- Data monitoring committees have been used in pharmaceutical industry clinical trials since at least 1988.
- Data monitoring committees can add to the objectivity and credibility of trials.
- Government-sponsored clinical trials are *research*-oriented; private industry sponsored trials are *development*-oriented.
- A single clinical trial cannot assess safety for rare events.
- Safety problems in drugs are often discovered postmarket.
- Data monitoring is needed for all phases of clinical trials.
- Issues arising in safety monitoring may differ by the size of the sponsoring pharmaceutical company.

1.1 What Is a Data Monitoring Committee (DMC)?

Professor Jerome Cornfield once defined a clinical trial as *cogent* description (Cornfield, 1973). It has become clear that objectivity in reviewing accumulating data in clinical trials is extremely important in maintaining cogency. One factor that can operate against cogency is *bias.* We will discuss the concept of bias in Chapter 6, but for now let us define it as a conscious or unconscious lack of objectivity due to a sponsor staff's interest in getting the experimental treatment approved by regulatory agencies. Clinical trial sponsor staff can introduce bias into trial conduct if they review efficacy data during the trial. They may have a tendency to underplay the importance of adverse events that present during the trial. This is especially true in oncology trials, which are usually conducted with sponsor staff, investigators, and patients all aware of treatment assignment. Although sponsors tended to claim that their trial management did not introduce bias, the current feeling is that if clinical trial results are to be persuasive to regulatory agencies, practicing physicians and the general public, even the appearance of bias must be avoided. The problem is complicated by the fact that

investigators and others associated with the trial can also introduce bias. Although bias reduction is important for scientific and regulatory reasons, it has also become evident that objective review of accumulating data is necessary to protect patient safety.

In the late 1980s, following the example of clinical trials run by the government through agencies such as the National Institutes of Health (NIH), the Veterans Administration (VA), the Center for Disease Control and Prevention (CDC), the British Medical Research Council, and the French Inserm, the pharmaceutical industry began to form DMCs. These committees took on various names and were of various forms. Initially some of these committees included members who were sponsor staff, but regulatory agencies took a dim view of this practice. Eventually membership evolved to individuals who were not members of the sponsor staff but were physicians of appropriate specialties and experienced clinical trial biostatisticians who could be trusted to review efficacy and safety data in such a way that bias would be minimized. In coming chapters we will review "best practices" for minimizing bias and for DMC operations in general. This book will concentrate on the safety role of the DMC in industry-sponsored trials. Safety data constitute 85% of data collected in clinical trials submitted to the U.S. Food and Drug Administration (FDA; Rochester, 2008), are evaluated more subjectively than efficacy data and, experience shows, constitute about 90% of DMC operations.

1.2 Some Definitions

It will be useful to provide some brief definitions of important terms. We will return to these terms to provide more rigorous definitions later in the book.

The *sponsor* of a trial is the organization that has the ultimate responsibility for reporting the results to the regulatory authorities. For our purposes it will most often be a pharmaceutical or biotechnology company, but it could be a university, government agency or, in the case of orphan drugs, a patient–parent support group.

An *adverse event* (AE) is any unfavorable and unintended sign (including an abnormal laboratory finding, for example), symptom, or disease temporally associated with the use of a drug, whether or not considered related to the drug. An adverse event is deemed *treatment-emergent* if the adverse event is not a manifestation of a condition that existed prior to the clinical trial. It is not always easy to make this distinction. A *serious* adverse event (SAE) is any untoward medical occurrence that, at any dose, results in death, is life-threatening, or requires inpatient hospitalization or prolongation of existing hospitalization. This is the regulatory definition. We will see in Chapter 4 that this definition may generate different types of adverse events in different countries due to differences in hospitalization policy.

The reader will be familiar with patients, and possibly investigators, being *blinded* to treatment assignment. In deference to our ophthalmology colleagues we will use the term *masked* as a synonym for blinded. When treatment assignment is not masked to anyone, we usually use the term *open label.*

The term *monitoring* is used in several different contexts in the pharmaceutical industry. *Statistical monitoring* refers to making calculations on accumulating efficacy data to justify early termination of a clinical trial (Lan, Proschan, and Wittes, 2006) and/or sample size reestimation (Chuang-Stein, Anderson, Gallo et al., 2006). *Safety monitoring* by sponsor staff and DMC members refers to continual review of accumulating safety data during the trial. *Site monitoring* is a quality control procedure applied periodically during the trial by sponsor or contract clinical research associates (CRAs; Woodin and Schneider, 2003).

1.3 DMC in Federal Government-Sponsored Clinical Trials versus Pharmaceutical Industry Clinical Trials

DMCs had long been included in federal government-sponsored clinical trials before they appeared in pharmaceutical industry clinical trials. The latter took on a different form than the former due to differences in the characteristics of the trials being conducted (Herson, 1993). Table 1.1 provides a summary of differences between federal government-sponsored trials and private industry-sponsored trials. The federal government trials are primarily *research* or *science* oriented with a public audience, are sometimes community based (see, for example, Djunaedi, Sommer, Pandji et al., 1988), and most often involve drugs already approved such as the Women's Health Initiative progestin trial (Writing Group for the Women's Health Initiative Investigators, 2002; Wittes, Barrett-Connor, Braunwald et al., 2007). Industry trials are patient-based and *development* - or *product*-oriented with the goal of convincing regulatory agencies that a new product should be approved and then reaching a market segment of patients through physicians. To illustrate this distinction consider the NIH-sponsored Lipids Research Clinics (LRC) trials (Lipids Research Clinics Program, 1984a, 1984b). At the LRC design stage researchers asked the question, "Does cholestyramine treatment to lower low density lipoprotein (LDL) for patients with hyperlipidemia have an effect on mortality and morbidity?" After 10 years of research a positive result was found. Then private industry was able to follow with development trials of fewer patients and shorter duration to show that lovastatin, for example, was effective in lowering LDL (Havel, Hunningshake, Illingworth et al., 1987). This conclusion was considered acceptable for approval because the NIH trial had established that lowering LDL had a positive effect on a clinically significant endpoint. When this pivotal trial began, the sponsor research staff "knew" the answer—that is, on the basis of the preliminary trials they had confidence that the construct they designed would result in a positive outcome.

DMCs in NIH-sponsored trials are usually involved in trial design, sample size requirements, data analysis methods, data quality, publications policy, investigator evaluation, and so on in addition to efficacy and safety review. These responsibilities become more complicated when, in addition to DMCs, the NIH trials include steering committees and endpoint committees. The duties of DMCs operating within private industry trials are narrower in scope.

TABLE 1.1: Characteristics of Clinical Trials Utilizing DMCs by Sponsorship

Characteristics	Federal Government Sponsored	Private Industry Sponsored
Purpose	Advance medical research	Product approval
Activity	Research	Development
Orientation	Science	Product
Sampling unit	Community or patient	Patient
Audience	Public	Regulatory agency
At design stage	Know the question	"Know" the answer
Approval status of study drugs	Often approved drugs	Premarket, experimental
Design and analysis methods	Freedom to be creative	Must adhere to regulatory agency requirements
Pace	Careful, deliberate	Aggressive
Data quality control	Each trial can establish its own standards	Must adhere to high standards of Good Clinical Practices
Financing	Federal budget or grants to universities	Corporate
Potential for conflict of interest	Lower than private industry sponsored	Higher than federal government sponsored
Type of trial	One trial, large number of patients, long duration	Several trials with small number of patients and short follow-up time

Lachin (2004) has indicated that there is more of a chance for conflict of interest in private industry trials than in government-sponsored trials. Pharmaceutical firms have learned that good science and objectivity are the best strategies for shortening the time to approval. Thus DMCs in the pharmaceutical industry evolved as "blue ribbon" panels for independent certification on issues such as adjudication of efficacy endpoints, conduct of planned interim analyses, and safety monitoring. DMCs in the pharmaceutical industry are a node in an aggressive drug development process leading to marketing. There are considerable financial consequences in the outcome of the trial and numerous opportunities for bias and/or conflict of interest.

1.4 Stewardship

Although sponsors retain DMCs to add to the objectivity and credibility of trials, DMC members can best fulfill their obligations to sponsors and patients by considering themselves responsible for the *stewardship* of the trial. This implies both the

preservation of credibility of the trial and the aegis of patient safety. How this stewardship can best be carried out will be covered in later chapters of this book. For now it is sufficient to note that DMC members must be proactive and consider themselves "officers" of the trial if they are to fulfill their responsibilities to the patients and sponsor.

1.5 Some Recent History

1.5.1 Development of DMCs in the Pharmaceutical Industry

One of the first known DMCs in pharmaceutical industry clinical trials was the cimetidine stress ulcer clinical trial in 1988–89 (Herson, Ognibene, Peura et al., 1992). This trial was conducted in intensive care units and was designed to compare cimetidine to placebo for prophylaxis of upper gastrointestinal bleeding due to stress (Martin, Booth, Karlstadt et al., 1993). The primary efficacy endpoint for this trial was prophylactic failure defined as the appearance of bright red blood and other bleeding-related outcomes. A supplementary definition of failure was *insufficient therapeutic effect* (ITE), which investigators could invoke at their discretion to remove a patient from the study if they feared the patient might begin to bleed. The sponsor decided to create a DMC for independent and masked certification of bleeding data and determination if ITE decisions were made according to usual clinical practice and safety monitoring including judgments on whether death was disease-related. The decision to involve a DMC in this trial came from the sponsor's experience on earlier clinical trials for this product where the possibility of bias in sponsor staff efficacy classifications raised credibility issues with FDA. The sponsor and consultants used some aspects of DMCs on NIH-sponsored trials to write the charter for their DMC.

Physicians chosen for this committee had expertise in gastrointestinal disease and emergency medicine. The data for DMC review were sent by the sponsor to a contract research organization (CRO) with treatment assignments coming from a manufacturing office of the sponsor rather than from those sponsor staff involved in the trial. All data processing for the DMC was performed by the CRO. When the trial ended the DMC presented the results to the sponsor. However, those sponsor staff members still evaluating safety were not informed of results to avoid introduction of bias into ongoing safety evaluations.

1.5.2 Guidances—FDA, NIH, and ICH

Since DMCs first appeared in pharmaceutical trials in the early 1990s, much has been written about the role of DMCs. The FDA guidance finalized in 2006 (U.S. Food and Drug Administration, 2006), and the International Conference on Harmonisation (ICH) mentions DMCs in their E-3 guideline on clinical study reports (International Conference on Harmonisation, 1995), the E-6 guideline on good clinical practices (International Conference on Harmonisation, 1996), and the E-9 guideline on statistical principles (International Conference on Harmonisation, 1998).

Hemmings and Day (2004) provide a good discussion of regulatory issues related to DMCs. Literature oriented toward NIH-sponsored trials would include the NIH guidelines (U.S. National Institutes of Health, 1998, 1999, 2000), the DAMOCLES literature search (Sydes, Spiegelhalter, Altman et al., 2004) and books by Ellenberg, Fleming, and DeMets (2002) and DeMets, Furberg, and Friedman (2006). Recently some prestigious medical journals have adopted a policy of not publishing results of industry-sponsored trials unless an independent DMC was involved (Fontanarosa, Flanagin, and DeAngelis, 2005).

1.5.3 Other Vehicles for Patient Protection

DMCs are not the only source of protection of patient safety. Each clinical trial site (hospital, clinic, doctor's office) comes under the auspices of an Institutional Review Board (IRB) or Ethics Committee which reviews protocols and their amendments and receives periodic safety reports. Each sponsor has internal safety review mechanisms and some larger medical institutions have internal DMC-like committees who review safety on ongoing trials regardless of sponsorship. The relationship between these entities and DMCs will be covered later in this book.

1.6 DMC's Place in the Drug Development Cycle

1.6.1 Phases of Drug Development

Drug development is often broken into several phases—molecular, preclinical, exploratory, confirmatory, and postmarket (Scheiner, 1997). The exploratory and confirmatory phases include those clinical trials that will be used for drug approval. The exploratory trials are often referred to as phases I and II. These trials are primarily *proof of concept* trials which establish dosage (maximum tolerated dose) and efficacy (minimum effective dose). Some prefer to call these trials *test of concept* rather than *proof*, but in any case it is important to consider these early trials as *learning trials*. These are followed by the confirmatory trial(s) known as phase III. Here we apply what we learned in earlier phases regarding dose, schedule, and appropriate patient populations to design a trial that is expected to demonstrate efficacy and safety with statistical precision. Phase III trials are sometimes called *pivotal* trials because they are the trials that will form the basis of the regulatory decision. The terms *confirmatory, phase III,* and *pivotal* will be used synonymously in this book.

1.6.2 Limitations of a Clinical Program for Revealing Safety Issues

The clinical program is conducted to produce evidence of efficacy and safety sufficient for marketing approval by regulatory agencies. This program will provide evidence of serious adverse events that occur with highest frequency. It must be understood that it is the DMC's role to consider these adverse events but, obviously, not to be responsible for all SAEs that may ever be associated with the drug during its lifetime.

TABLE 1.2: Clinical Trials Terminated Due to Safety

Drug	Indication in Trial	Risk	Trial Terminated
torcetrapib	Raise HDL (high density lipoprotein)	Increase in cardiovascular events and death	2006
estrogen + progestin	Prevention of chronic disease	Invasive breast cancer, coronary heart disease	2002
tirilazad	Head trauma	Death	1994
rofecoxib	Polyp prevention	Thrombotic events	2004
celecoxib	Polyp prevention	Thrombotic events	2004
naproxen	Alzheimer's	Fear of thrombotic events	2004

A single pivotal clinical trial designed to demonstrate efficacy will not be able to assess rare events or those that represent delayed effects. The DMC may not gain an understanding of which AEs will become chronic during the duration of a clinical trial. Despite the limitations, several clinical trials have been terminated for safety in recent years. A partial listing is found in Table 1.2.

In a clinical program it is always useful to keep in mind the "rule of 3000/n" where n equals the number of patients exposed to the drug in a clinical program. If $n = 1000$ patients, then the clinical program is likely to find at least one case of AEs that occur at an incidence of $3{,}000/1{,}000 = 3/1{,}000$. If $n = 500$, then the clinical program would be sensitive to find at least one case that occurred with incidence 6/1,000. To be able to find an adverse event that occurs at the rate of 1/100,000 a program of 300,000 patients would be required. This type of AE could be found only in postmarket surveillance. Table 1.3 presents a table for sensitivity for AE detection in a hypothetical clinical program for development of a diabetes drug. We see the sensitivity of the individual

TABLE 1.3: Sensitivity to AE Detection in a Clinical Program for Development of a Diabetes Drug

Protocol No.	Description	*n* for this Protocol	AE Detection Rate/1000	N (Cumulative)	Cumulative AE Detection Rate/1000
11	26 wks + 3 mo extension	348	8.6	348	8.6
12	26 wks + 3 mo extension	406	7.4	754	4.0
21	26 wks + 6 mo extension	510	5.9	1264	2.4
22	26 wks + 6 mo extension	928	3.2	2192	1.4
31	52 weeks	604	5.0	2796	1.1

protocols to detect AEs ranges from 3.2 to 8.6/1,000. As the program progresses sensitivity does not drop below 3/1,000 until the first 26 week + 6 mo extension trial (protocol 21). At the conclusion of the program with 2796 patients cumulative AE sensitivity is 1.1/1,000. There are many potential AE types that would occur at a rate of 1/100,000 and thus the need for postmarket surveillance to clarify the safety profile of the drug.

Table A.1 in the Appendix presents a table of adverse events observed in placebo-controlled trials for marketed drugs together with the number of patients enrolled on each treatment arm. The list reveals that the adverse events observed are the most common and not necessarily most serious. The postmarket phase will reveal the rare and potentially serious adverse events. We will revisit this table again in various chapters.

Lin, Chern, and Chu (2003) acknowledge this limitation but are concerned that failure to find cases of liver toxicity in a confirmatory trial might lead to a conclusion that the experimental drug is not associated with liver toxicity. They provide useful guidelines for surrogates, such as laboratory values, to liver disease that might be uncovered in a clinical trial. DMC members can, presumably, discuss surrogates for other diseases when appropriate.

Postmarket experience will be needed to uncover SAEs that occur with low incidence. The clinical programs generally enroll patients with much narrower characteristics than those who will receive the drug after approval. Eligibility requirements specify strict age groups and prohibit enrollment of patients with certain medical histories, comorbidities, and concurrent medications. As larger numbers and newer types of patients are exposed to the drug postapproval, newer SAEs are likely to emerge. In the postapproval era papers appear in the literature presenting safety profiles of drugs used over many controlled clinical trials. Examples would include Schoenfeld (1999) for gastrointestinal safety of the nonsteroidal anti-inflammatory drug meloxicam; Strampel, Emkey, and Civitelli (2007) for the safety profile of bisphosphonates in the treatment of osteoporosis; and Wernicke, Lledo, Raskin et al. (2007) for the cardiovascular safety profile of duloxetine used to treat major depressive disorder. A summary of limitations of a clinical program to uncover important safety issues is presented in Table 1.4.

TABLE 1.4: Limitations of Safety Assessment by a DMC in a Single Clinical Trial

1. DMC is likely to find only AEs that occur immediately and with highest frequency.
2. DMC is not likely to find rare or delayed effects.
3. DMC may not develop an understanding of chronic effects.
4. Due to stringent eligibility requirements, clinical trial patients are not representative of those who will be treated with the drug after approval. Different patient types on varying concomitant medications may have a different but more common safety experience than the clinical trial patients.
5. DMC may miss subtle signals that involve extensive analysis and additional data on surrogates of AEs.

1.6.3 Postmarket Safety Actions

After a drug is approved by the FDA, there are several actions the agency may take when safety concerns arise. These actions include ordering the drug withdrawn from the market, attaching a *black box warning* to the drug's package insert (i.e., label change to highlight the description of the SAE in the package insert) and ordering discontinuation of a dosage form. Carpenter, Zucker, and Avorn (2008) report that during the period of 1993–2004 the FDA made 11 safety-based withdrawals and 14 black box warnings (21 drugs had either a withdrawal or black box warning or both) and 36 had dosage form discontinuation. Lasser, Allen, Woolhandler et al. (2002) indicate that there were 548 new chemical entities approved by FDA during the period of 1975–1999, and of these, 56 drugs (10.2%) acquired a black box warning or were withdrawn. On the basis of the FDA's adverse experience reporting system during the period of 1969–2002, a total of 75 drugs or drug products were removed and 11 received restricted prescription requirements (Wysowski and Swartz, 2005). Table 1.5 presents a partial list of drugs withdrawn from the market for safety reasons during the period of 1975–2007.

Lasser, Allen, Woolhandler et al. (2002) list 28 drugs for which black box warnings were issued during the period of 1975–2000. The timing of the warnings ranged from 1 to 23 years postapproval. The frequency of black box warnings appears to have accelerated since 2000 with adverse events of new awareness and interest and new sources of safety evidence. For example, since 2000 fifty drugs received black box warnings for suicide risk (Mundy, 2008). Rosiglitazone, indicated for diabetes, received a black box warning for risk of cardiovascular risk with much of the evidence coming from a meta-analysis published by an academic cardiologist (Harris, 2007; Nissen and Wolski, 2007).

There is considerable controversy about the timing and appropriateness of specific postapproval drug actions (Lasser, Allen, Woolhandler et al., 2002; Carpenter, Zucker, and Avorn, 2008; Lurie and Sasich, 1999; Friedman, Woodcock, Lumpkin et al., 1999). Much of this controversy stems from a misunderstanding of the extent that a clinical program can reveal safety concerns and the differences of opinion in the ability and methodology to assess early safety signals from clinical trials. The following chapters will explain how DMCs can help in identification of safety signals.

1.6.4 Role of DMCs in Exploratory and Confirmatory Trials

This book will concentrate on DMCs in confirmatory trials where they are used most frequently. There is no doubt that there is a need for safety monitoring in exploratory trials especially because these represent the first use of new drugs in humans. However, it is not clear that this must be accomplished through a completely independent DMC, and many feel that engaging a completely independent committee at this stage would slow down the development process. In a paper commissioned by the Society for Clinical Trials, Dixon, Freedman, Herson et al. (2006) give useful guidance on this issue as do Hibberd and Weiner (2004).

The phase I trial uses objective safety data and the protocol team in house together with participating investigators can generally handle the safety monitoring with little question of credibility. Phase II does not usually require a completely

TABLE 1.5: Partial List of Drugs Withdrawn from Market for Safety Reasons, 1977–2007

Drug	Indication/Class	Risks	Approved	Withdrawn
aprotinin	Reduce blood loss during cardiovascular surgery	Complications of surgery, death	1998	2007
pergolide	Parkinson's disease	Heart valve damage	1988	2007
tegaserod maleate	Irritable bowel syndrome	Myocardial infarction, stroke	2002	2008
valdecoxib	Pain, anti-inflammatory	Heart attack, stroke	2001	2005
rofecoxib	Pain	Thrombotic events	1999	2004
cerivastatin	Lipid lowering	Muscle damage	1997	2001
rapacuronium bromide	Injectable anesthetic	Bronchospasm	1999	2001
alosetron	Irritable bowel syndrome in women	Intestinal damage	2000	2000
cisapride	Night heartburn	Fatal heart rhythm	1993	2000
troglitazone	Type 2 diabetes	Severe liver toxicity	1997	2000
astemizole	Antihistamine	Fatal heart rhythm	1988	1999
grepafoxacin	Antibiotic	Fatal heart rhythm	1997	1999
mibefradil	High blood pressure and chronic stable angina	Dangerous interactions with other drugs	1997	1998
bromfenac	Pain	Severe liver damage	1997	1998
terfenadine	Antihistamine	Fatal heart rhythm	1985	1998
fenfluramine	Obesity	Heart valve abnormalities	1973	1997
dexfenfluramine	Obesity	Heart valve abnormalities	1996	1997
flosequinan	Cardiovascular disease	Increased mortality	1992	1993
temafloxacin	Antibiotic	Hemolytic anemia, renal failure, etc.	1992	1992
encainide	Antiarrhythmic	Increased mortality	1986	1991

(*Continued*)

TABLE 1.5: Partial List of Drugs Withdrawn from Market for Safety Reasons, 1977–2007 (Continued)

Drug	Indication/Class	Risks	Approved	Withdrawn
nomifensine	Antidepressant	Hemolytic anemia	1984	1986
suprofen	Analgesic, NSAID[a]	Pain	1984	1986
zomepirac	Analgesic, NSAID[a]	Anaphylaxis	1980	1983
benoxaprofen	Analgesic, NSAID[a]	Jaundice	1982	1982
ticrynafen	Antihypertensive	Hepatic toxicity	1979	1980
azarbine	Psoriasis	Thromboembolism	1975	1977

[a] Nonsteroidal anti-inflammatory drug.

Sources: U.S. Food and Drug Administration (2002), Safety-based drug withdrawals (1997–2001), http://www.fda.gov/FDAC/features/2002/chrtWithdrawals.htm; Lasser, K.E., Allen, P.D., Woolhandler, S.J. et al. (2002) Timing of new black box warnings and withdrawals for prescription medications, *JAMA*, 287, 2215–2220; FOI Services (2008) document search, http://www.foiservices.com.

independent DMC, but it often makes sense to include one or two outside members (physician or physician plus biostatistician) to the protocol team with the understanding that one of the outside people will take the role of chair of the DMC when the drug enters confirmatory trials. This allows for some outside expertise in the first efficacy trials and provides drug familiarity for the confirmatory DMC. Of course if the trial is first for a novel drug such as a drug eluding stent or one utilizing gene therapy or nanotechnology, it may be advisable to add to the DMC additional outside reviewers with the particular expertise.

1.6.5 Investigator-Sponsored Trials

In the U.S., investigator-sponsored trials are those conducted by an academic physician (investigator) using drugs provided by a pharmaceutical company. The investigator, rather than the sponsor, is responsible for all interaction with the FDA. Thus, sponsors have little control over investigator-sponsored trials, but there is clearly a need for safety review. When the investigator-sponsored trial is conducted within a single institution and that institution has a standing internal DMC to monitor trials that do not otherwise have a DMC, this body would usually be sufficient. If no such panel exists or the trial is multicenter, investigators might want to consider some of the ideas for exploratory trials above.

1.6.6 Open Label Extension Studies

For chronic conditions such as epilepsy, Parkinson's disease, and hypertension, it is common for patients who exit phase II or III trials to be put on open label extension studies. These trials are uncontrolled and have the purpose of obtaining more precision in estimation of incidence of adverse events and, perhaps, to uncover new adverse

events encountered in long-term exposure. DMCs may be asked to review data from ongoing open label extension studies while they are reviewing data from controlled trials. At the very least the DMC should review the extension study results at the end of the final confirmatory trial. Day and Williams (2007) provide insight into the role of the open label extension study in drug development.

1.7 Pharmaceutical Industry Demographics

1.7.1 Size of Companies

For the purposes of this book, the global pharmaceutical/biotechnology industry will be divided into three gross size groups by annual revenues in 2006. The term *Big Pharma* will apply to those companies with annual revenues greater than $8 billion. Companies with annual revenues of less than $8 billion but with products on the market will be called *Middle Pharma*. Those companies working in the development of their first product will be called *Infant Pharma*.

1.7.2 Public versus Private Companies

All of the Big Pharma companies and most of the Middle Pharma companies are publicly owned. Many of the Infant Pharma firms are publicly owned and those private ones are financed by venture capital firms seeking to raise additional rounds of financing while pursuing an exit strategy which would consist of taking the company public or selling it to a larger company. Today's reality is that, regardless of size, these companies are vulnerable to information that must be reported to investors on company activities and especially R & D (research and development) activities. Actions taken by DMCs can affect the financial status of these companies although differently depending on size. We will return to this important topic at various times in this book. For purposes of contrast we will be referring to Big versus Infant Pharma in much of what follows. As would be expected, Middle Pharma shares some characteristics of both its big and little brothers. Middle Pharma is very dependent on Big Pharma as a marketing partner for its products and for investment in R&D programs. Middle firms are very dependent on public markets for financing their new products. If a Middle firm has not had a product approved since it emerged from Infancy and has had a string of disappointments since emerging, it would be highly vulnerable to DMC negative decisions. The differences in the three levels of companies are further described in Table 1.6.

1.8 Conclusion

We have now seen the rationale and the setting for DMCs in the pharmaceutical industry. In the next chapter we will learn more about the members of the safety monitoring team and their roles.

TABLE 1.6: Characteristics of Pharmaceutical/Biotechnology Companies by Size

Characteristic	Big Pharma	Middle Pharma	Infant Pharma
Annual Revenues ($ in 2006)	More than $8 billion	Less than $8 billion	None from product sales
Products on Market	Many products on market	Some products on market	No products on market
Financial organization	Public	Public	Public or venture capital financed
Corporate goal	Expand product line	Expand product line, create corporate partnerships with Big Pharma	Show progress to increase financing, license the product to larger company
Clinical program financing	Complete	Almost complete	Trial by trial basis, further funding for pivotal trial might depend on results of interim analysis
Vulnerability to negative trial information	Small, unless drug is successor to blockbuster	Modest, especially vulnerable if it has been some time since last product approved	Considerable
DMC procedures	In place	In place	Not well developed, often created as trial develops

DMCounselor

Q1.1 I agreed to serve on a DMC for an Infant Pharma company which has gone public. The drug is a novel approach to pancreatic cancer. The trial is actually phase II but their regulatory consultant feels that if the results are positive the regulatory agency will consider it phase III. The sponsor has now told us that we will meet only once at the end of the trial because their board of directors is concerned that if we recommend early trial termination due to a safety concern, they would have to include this information in a press release and this would have a bad effect on their stock price. I would like to walk away from this DMC but this is an important drug and I would like to be associated with its development. What should I do?

A I had doubts about this sponsor when I heard that they have a consultant who told them that a phase II trial would count as a phase III trial in this case. It is doubtful that sufficient safety data would arise in a single phase II trial. In any case this sponsor's restrictions do not allow the DMC to fill the stewardship role. It is the DMC's responsibility to decide how often they will meet, not the sponsor's. The DMC must review accumulated data during the trial so that patients are not put at risk for a trial's duration if serious concerns arise during the trial. If this sponsor is afraid of an interim recommendation, why would they want a DMC to make an assessment at the end of the trial? You should try to convince the sponsor not to begin the phase II/III until they have more confidence in the safety of this product. At that point there should be much less financial risk in having a DMC schedule periodic meetings during the trial.

Q1.2 I was asked to serve on a DMC for a phase II trial. I said that I would do so provided I would automatically be placed on the DMC for the phase III trial. The sponsor refused. Why would they do this?

A The sponsor was right in this instance. Although it is a good idea to have some continuity between phase II and phase III, it is also good to have some new people on the phase III trials. In choosing which, if any, phase II DMC members would carry over to the phase III DMC, sponsors would usually wait until the phase II trial was concluded, at which time they would have learned from that trial what type of expertise would be needed for the phase III DMC.

Q1.3 I was asked to serve as a biostatistician member of a DMC. I accepted and found out that the trial was phase I. Is a biostatistician member really needed for a phase I trial?

A Outside members are not usually employed on phase I trials but the sponsor appears to think it is necessary in this case. Do not think that you are not needed just because you will not be looking at confidence intervals and explaining survival curves to the physician members. Your knowledge of protocols, objectivity, logical thinking, and so on would be very important to the committee.

Q1.4 I was asked to be a physician member of a DMC for a neurology drug by a Middle Pharma company. I accepted but later found out that the trial was completed 5 months ago without a DMC. The sponsor is now in negotiations to license the drug to a Big Pharma company and the latter insists that there be an independent review of safety before talks can continue. Is this an appropriate use of my time?

A The committee the sponsor is forming is not a data *monitoring* committee but an ad hoc committee to come in once and make statements about safety presumably by also taking

efficacy into account. This committee will obviously be unmasked and will not have had the benefit of considering and scrutinizing safety issues as they arise. If you feel you want to serve, this is OK, but make sure that the sponsor does not represent your committee as a DMC. The sponsor must also understand that your committee is not coming together for two hours to be a "rubber stamp" on safety. If your committee needs more time and information, it must be granted. Also your committee should not be brought to the table either as individuals or collectively to be part of business negotiations with the company purchasing the license.

Chapter 2

Organization of a Safety Monitoring Program for a Confirmatory Trial

Bullets to Remember

- A DMC is a committee independent of the sponsor.
- DMCs typically consist of both physicians and biostatisticians.
- Sponsors should have standard operating procedures for the creation, organization and management of DMCs.
- Even the appearance of conflict of interest among DMC members must be avoided.
- Both the sponsor staff and the DMC members are "officers" of the clinical trial and must respect one another's authority and responsibility.

2.1 Members of the Safety Monitoring Team

2.1.1 The Sponsor

As was indicated earlier, the *sponsor* is the organization that pays for the trial and has the objective of getting the experimental drug approved by regulatory agencies. The sponsor will have the task of appointing a DMC before the start of the clinical trial and ensuring its independence. An important person within the sponsor organization is the *sponsor representative* to the DMC. This person is typically the senior clinical research professional in charge of the trial. This person will coordinate sponsor activities and those of the other elements of the safety monitoring operation. The person will generally be in charge of a protocol team of people who will work with the DMC. These team members will generally represent clinical operations, biostatistics, data management, and safety surveillance. The sponsor representative might possess an MD, PhD, or PharmD. However, regardless of degree, this individual will have the clinical research experience necessary to be leading the trial. In Infant Pharma the sponsor representative might be the vice president of clinical research or the medical director of the company, or it might be a person with extensive clinical operations

TABLE 2.1: Desirable Characteristics in a Sponsor Representative to the DMC

1. Respect the DMC as being responsible for the stewardship of the trial and remember that the DMC members are at least partially unmasked to treatment and sponsor staff is completely masked
2. Respect the independence of the DMC
3. Interact and respect the views of the entire committee not just the chair or the person of a certain discipline
4. Respect the views of other sponsor staff especially with their interactions with DMC members of same discipline
5. Be open-minded about DMC requests for ad hoc computer programming, consultants, etc., but not intimidated about asking for justification
6. Be knowledgeable about clinical trial operations in general and at the CRO and DAC that may be involved in this trial
7. Be knowledgeable of sponsor, standard operating procedures (SOPs) for DMCs, the DMC Charter, and FDA and ICH guidelines for DMCs and safety reporting
8. Be knowledgeable about the clinical issues in the disease being treated, experimental drug, and its safety profile and clinical trial history
9. Be knowledgeable of statistics as applied to clinical trials

experience because the person acting as medical director is a contract consultant to the sponsor. In Big Pharma and Middle Pharma the sponsor representative will mention several people above him or her in the organization who have responsibility for the therapeutic area and the trial at hand. It will be important for the DMC members and especially the chair to understand who the DMC reports to under different circumstances.

Many characteristics for a sponsor representative would be considered desirable. The most important is that the sponsor recognizes the DMC as being responsible for the stewardship of the trial and being independent of the sponsor. Table 2.1 presents a list of desirable characteristics for a sponsor representative. The rationale for many items on this list will become more obvious in later chapters of this book.

2.1.2 Data Monitoring Committee

The DMC is the subject of this book. Much more will be said about its composition and functions later. For now it is sufficient to reinforce that the DMC is an *independent* committee of individuals, with credentials in medicine, biostatistics, and so on, who are not employees of the sponsor or investigators/biostatisticians on the trial having the responsibility of stewardship for the trial. In this capacity the committee will protect patient safety by periodic review of safety data and responding to trends in serious adverse events whenever they may occur. In some trials DMCs have efficacy monitoring responsibilities as well, but, apart from risk-benefit analyses to be discussed later, DMC efficacy responsibility is beyond the scope of this book.

2.1.3 Data Analysis Center

The Data Analysis Center (DAC) is the organization that will prepare tables and reports for the DMC under formats requested by the DMC. The DAC could be the statistical unit of the sponsor or a CRO. A biostatistician from the DAC will be called the *independent statistician* and will be a nonvoting ex officio member of the DMC. The DAC will have the treatment assignment codes, and thus, members can unmask themselves at any time. It is thus preferable that the DAC not be the same organization working on the ultimate regulatory submission. However, sponsors and CROs working on the regulatory submissions have found satisfactory ways of building a "firewall" between those providing DAC services and those working on the regulatory submissions. Much has been written about the need for independence or ways of constructing this firewall (see for example Ellenberg and George, 2004; Siegel, O'Neil, Temple et al., 2004; Snapinn, Cook, Shapiro et al., 2004). Nevertheless the degree of independence that actually exists in any situation remains a controversial topic. More will be said later about the DAC and its place in data flow.

In addition to statistical analysis DACs sometimes perform administrative functions such as making travel arrangements and paying DMC members. However, DACs should be chosen for their ability to support safety surveillance in an ongoing clinical trial, not primarily on the basis of computer programming or administrative abilities. A list of desirable characteristics for a DAC is found in Table 2.2.

In some clinical trials, especially those run in Europe, the biostatistician member of the DMC is given a safety data set prior to the DMC meeting, and this individual prepares the data for review and performs all other functions that the DAC biostatistician would normally perform. There is no problem with this approach but in what follows we will assume that the DMC is working with a separate DAC biostatistician. In trials where the DAC biostatistician has responsibility for preparing tables

TABLE 2.2: Desirable Characteristics for a Data Analysis Center

1. Experience in serving as a DAC
2. SOPs for DAC operations
3. SOPs for software validation
4. Knowledge of FDA and ICH guidelines for safety reporting and Data Monitoring Committees
5. Knowledge of MedDRA (Medical Dictionary for Regulatory Affairs; see Chapter 4) coding of AEs
6. Library of validated software for report generation and statistical analysis
7. Flexible staff for timely response to ad hoc requests
8. If needed by sponsor, support of administrative services such as travel and meeting arrangements, host conference calls
9. Off-site computer backup
10. Statistical staff with knowledge of
    a. Clinical trial statistical methods
    b. Interim analysis methodology including conditional power and predictive power (these terms will be defined in Chapter 5)

and doing statistical analysis, it is important to decide, at the outset, if the DMC biostatistician member can also receive data to do additional analysis, and the cost control issues and procedures of such a practice must also be discussed before the trial begins.

2.1.4 Institutional Review Board

An Institutional Review Board (IRB) exists at every institution performing medical or behavioral research on human subjects in the United States. IRBs are regulated by the FDA and the Office for Human Research Protections in the Department of Health and Human Services (Code of Federal Regulations, 2005). In other parts of the world similar committees are often referred to as ethics committees. The IRBs are responsible for reviewing proposed and ongoing research at their institutions to decide if the research is ethical and deciding of informed consent is sufficient and appropriate safeguards, such as the existence of a DMC, are in place. IRBs review protocols and their amendments, investigator brochures and amendments, serious adverse event reports, and so on. There is some overlap with DMCs, but the DMC is responsible for the stewardship of the trial over all institutions whereas the IRBs have responsibility within their institutions. The DMC reviews all trial safety data on a regular basis, can be unmasked but receive at least partially unmasked data and can recommend termination of the trial. The IRB reviews only trialwide serious adverse events that are serious, drug-related and unexpected. The IRB also reviews an annual report of SAEs judged to be related to drugs and can recommend termination of the trial at their institution.

We will see that DMCs will review more information than this but they should certainly review no less. A summary of IRB and DMC roles in safety surveillance is presented in Table 2.3.

2.1.5 Scope of DMC Authority

DMC opinions on safety are advisory to the sponsor. It should not be assumed that if sponsor and DMC are in agreement, there are no serious safety issues and the drug is now automatically classified as safe. The regulatory agencies will have the final say in this matter.

2.2 How Is a DMC Created?

The sponsor has the responsibility for creating the DMC. This is generally done as soon as most investigators are selected and it is clear that certain physicians who could serve as investigators will not be serving. This latter group would serve as a pool for selection of physician DMC members. Ideally the DMC should be in place before the first patient is randomized to the trial. Unfortunately this is not always the case

TABLE 2.3: Safety Monitoring by Institutional Review Boards versus Data Monitoring Committees

Characteristics	IRB	DMC
Origin	Required by law in most countries	Recommended by regulatory agencies for phase III clinical trials
Purpose	Big picture—ethics and safety, at institutional level	Detailed safety review trialwide
Review expedited adverse events (serious, possibly related, unexpected)	Yes	Yes
Other adverse events	Annual update of related	Frequent update of all adverse events
Treatment group information	No	Yes
Review: protocol	Yes	Yes
Investigator brochure	Yes	Yes
Informed consent	Review and approve	Review but not approve

(see *DMCounselor* Q2.2 below). Typically, the sponsor's study team would meet to go down a list of candidates, prioritize them, and then begin contacting candidates from the top of the list.

All activities for the creation, organization, and management of DMCs should be covered in the sponsor's standard operating procedures (SOPs). These SOPs would typically be found as part of SOPs for the sponsor's overall risk management plan (Haas, 2004; Bush, Dai, Dieck et al., 2005). More will be said of SOPs for DMCs later, but it is important to note here that, like many drug development activities, sponsor staff should not be forming a DMC by intuition but rather by following documented procedures. The need for DMC SOPs applies to all sponsors regardless of size. However, experience shows that Big Pharma and Middle Pharma have SOPs in place and most of their clinical staff members are well versed in, or at least familiar with, these procedures. Infant Pharma companies are often creating procedures as they go along, partly because of time constraints and partly because of the need to create a balance between differing procedures that staff bring with them from their former companies. This can sometimes be frustrating for DMC members, but it should be looked at as an opportunity to help the startup group create the best procedures.

The principal organizational document for the DMC is the DMC Charter. The charter is an outgrowth of the SOPs and indicates the responsibilities of all parties within and outside the DMC. More will be said about the DMC Charter in the next chapter.

2.3 Membership

Let us now consider some of the characteristics appropriate and not appropriate for DMC members.

2.3.1 Physicians

The physician members would be of two types—those with expertise in the indication under investigation and those with expertise in expected adverse events. In a trial for rheumatoid arthritis we would certainly want to include rheumatologists who specialize in rheumatoid arthritis. However, if it is known that cardiovascular events (moderate hypertension, transient arrhythmia) are likely, a cardiologist might be included. For a drug used to treat diabetic shock a diabetes specialist would be needed but also, perhaps, an emergency medicine expert. In multinational trials it is advisable to include physicians who practice in the various cultures. These members can fill the usual physician role as well as advise on cultural issues that might affect the nature of adverse event reports. More will be said about DMC issues in multinational trials in Chapter 4.

2.3.2 Biostatisticians

A biostatistician will be needed on the DMC. This person should be experienced in the indication for the trial and familiar with statistical methods for safety analysis as well as efficacy analysis. If innovative methods of design and analysis are to be employed (e.g., adaptive designs, Bayesian methods), the biostatistical member should be well versed in these techniques and be able to explain them to the physician members.

2.3.3 How Many Members Are Needed?

Although having three members—two physicians and a biostatistician—appears to be common, the precise number needed depends on the complexity of the trial and the various kinds of expertise needed. Suffice it to say that the number of members should be the minimum needed to cover the waterfront of expertise. DMCs with more than one biostatistician are rare. As mentioned above the DAC will contribute their own biostatistician to the committee as a nonvoting member. As we will see later, scheduling DMC meetings is not easy even when they are done as telephone conference calls. The more people, the more difficult it is to schedule meetings. If there were a five-person committee on an oncology trial with one of the members being a neurologist as an AE expert (i.e., neurological AEs expected), and three members present constituting a quorum, it is possible that the trial could run to completion with the neurologist never attending a meeting and this input would be marginalized. Of course meetings should be scheduled so that no member misses a meeting. The quorum rule should be invoked only if a member cannot attend due to a last-minute emergency. Even then the DMC chair can get input from the absent party over the next few days following the meeting.

2.3.4 Ad Hoc Consultants

One way to keep a DMC at reasonable size is to add ad hoc consultants when issues arise that require expertise not present on the committee. For example, it might make more sense to bring an allergist in to consult on hypersensitivity events than to have this person as a sitting member of the committee. Although ethicists and patient advocates are often present as DMC members on trials sponsored by the NIH (Friedman and DeMets, 1981), they would usually appear as consultants in pharmaceutical trials except for those trials where serious ethical issues would occur regularly. An example of the latter in the field of psychiatric drugs would be in mood disorders (Charney, Nemeroff, Lewis et al., 2002). Recent examples in medical devices would include vagus nerve stimulators for depression (Schuchman, 2007; Rush, Marangell, Sackheim et al., 2005) and extracorporeal liver assist devices (Ellis, Hughes, Wendon et al., 1996). Sponsors generally rely on IRBs for input on the "big picture" ethical issues.

2.3.5 Ubiquitous DMC Members

It is best not to appoint people to a DMC who are currently serving on multiple DMCs, say, five or more. Often-used members may confuse issues and protocols among the various trials, and it is good to have many different people serving on DMCs instead of creating a power elite. Service on a DMC is confidential so sponsors must rely on the judgment of potential DMC members in this regard.

It is reasonable for those who have not served on DMCs to be unsure about serving when first asked. Table 2.4 presents a list of useful questions that potential DMC members might ask a sponsor before agreeing to serve on a DMC. Those using this table should beware of vague answers or promises to find answers that are not fulfilled in a timely manner. As a DMC member you will be on a committee responsible for the stewardship of the trial, similar to being a board of directors of the trial. In speaking to the sponsor representative it is important to ascertain if the sponsor sees the DMC with this responsibility or as a necessary appendage similar to certain clauses in an informed consent form. Any evidence of the latter should be treated with caution.

Similarly a potential DMC member may be unsure if he/she is the right person to be serving on a DMC. Table 2.5 presents questions that the potential DMC member might ponder to help decide if DMC service is right for him/her at least at this time. This table might also serve as a checklist for sponsors interviewing potential DMC members.

2.3.6 Disclosure of DMC Membership

The protocol for the trial should indicate the existence of a DMC but there does not appear to be any advantage to disclosing the names of the members of a DMC to investigators or the public at large until the trial is completed. If the identity of the members is known during the trial investigation, competitors or others might attempt to extract information about trial progress from them in informal settings.

TABLE 2.4: Questions to Ask Sponsor before Agreeing to Serve on a Data Monitoring Committee

1. May I first look over clinical trial protocol and draft DMC charter?
2. Who will the other members be either by name or by specialization?
3. Will contract include indemnification?
4. When does the trial begin?
5. How many meetings are proposed per year? How many are face-to-face?
6. What type of adverse events are expected; of what severity?
7. Will the DMC be responsible for efficacy or just safety?
8. Will the DMC be asked to comment on manuscripts reporting trial results before submission to a journal?
9. If asked to chair the DMC
   a. Is there a budget for ad hoc requests to the DAC?
   b. To whom in sponsor management do I report?
   c. Will the DMC have the freedom to design tables for review of safety or must the committee use tables already designed by the sponsor?
   d. If Internet interactive software is to be used by DMC members, how will training be handled? What are the resources for support/troubleshooting?
   e. Who are the other members of the DMC and by what process were they selected?
   f. In what countries will the trial be conducted?
   g. Will there be a central clinical laboratory for all centers?
   h. On what criteria was the DAC selected? What are their capabilities?

TABLE 2.5: Questions for an Individual to Ponder before Agreeing to Serve on a Data Monitoring Committee

1. Is my primary interest protecting patient safety (+) or earning a consulting fee (−)?
2. Do I have the time to devote to DMC service including flexibility for ad hoc emergency meetings (+)?
3. Do I like working on committees (+)?
4. Do I appreciate the interdisciplinary and multicultural nature of clinical trials (+) or do I tend to feel that people would agree with me if they only had the same training and background that I have (−)?
5. Do I feel that even though the sponsor is paying I that one I can remain independent of the sponsor (+)?
6. If asked to chair a DMC
   a. Am I a consensus builder (+) or do I feel that as chair I have the final say (−)?
   b. Do I have the time to get the committee together for ad hoc meetings, review ad hoc meeting material and to continually interact with sponsor staff on DMC matters (+)?

Note: + indicates favorable trait; − indicates negative trait.

2.3.7 Multiple Sponsorship

In diseases such as cancer, HIV, malaria, and so on, it is not unusual for the pharmaceutical industry sponsor to have cosponsors such as the National Cancer Institute or an other NIH component, World Health Organization, Gates Foundation, Cooperative Oncology Groups, and so on. These organizations may have existing committees that have functions similar to pharmaceutical industry DMCs and feel that these committees are sufficient to meet DMC requirements. Sponsors should make sure that all of their SOPs and regulatory requirements will be achieved by the existing committee (see *DMCounselor* Q2.3).

2.3.8 From Where Are DMC Members Recruited?

Typically DMC members come from the ranks of academia, government, nonprofit organizations, and semi- or fully retired professionals. People employed in the pharmaceutical industry or CROs would be eligible if they have no conflicts of interest but such persons rarely appear on DMCs because their degree of conflict of interest may change during the course of a trial.

2.4 Term

The usual term for DMC members is one trial or two trials if the two are being run simultaneously to satisfy regulatory requirements of submitting "at least one well-controlled trial" for marketing approval. Although there is a definite advantage to retaining one clinical person through several trials—phase II, III, IIIB—many sponsors avoid creating a "Supreme Court" where the same DMC serves throughout the lifetime of the product because of the advantages of getting different points of view and the danger that some members may make decisions on drug safety in early trials and not pay proper attention to new information in confirmatory trials. However, there is nothing wrong in an ethical or regulatory sense with the same people serving throughout.

2.5 Conflicts of Interest

Justice Potter Stewart once said, "I can't define pornography but I know it when I see it" (U.S. Supreme Court, 1964). Similarly, conflict of interest for DMC members cannot be precisely defined but, as was said in Chapter 1, persuasiveness of results requires that even the appearance of conflict of interest must be avoided. Sponsor SOPs should provide conflict of interest guidelines. These guidelines should require financial disclosure of equity interest in the sponsor or those with competing products, consulting

and/or investigator arrangements with the sponsor or competitors, proprietary interest in the drug under investigation or competing products. For some products, such as orphan drugs, there may be only one or two physicians with appropriate expertise to serve on a DMC. In these cases joint service must be allowed under the condition that the sponsor will not use the opportunity to find out information about the competitor's clinical program and that the DMC member will not divulge information about the competitor.

There will always be debate about whether the DMC is truly independent of the sponsor but following some reasonable steps should ensure that its members are at least more independent of the outcome of the trial than the sponsor.

Sponsor SOPs should require some level of financial disclosure prior to finalizing appointment to a DMC. DMC members should not have sizable consulting contracts with the sponsor at the time of DMC service. Sponsor SOPs should include a cutoff in dollars as to what is sizable. This cutoff would be an advantage to Infant Pharma because they will not have to be in a bidding war with Big Pharma for available talent.

2.6 Compensation

The level of compensation that DMC members receive for their service is closely related to conflict of interest. There is the perception that if DMC members were paid too well, they would be less independent of the sponsor. Sponsors should have guidelines for reasonable compensation, and this matter can be discussed among sponsors. In multinational trials the DMC will usually consist of members from different countries and the definition of reasonable compensation will differ among countries. As long as compensation levels are not extravagant Infant Pharma can afford to retain the same professionals for DMC service as Big Pharma. Of course compensation should not be so low that members cannot justify taking the time away from their day jobs to do the necessary work. DMC members should think of this activity as a service to drug development and not as a lucrative consulting sideline.

2.7 Liability and Indemnification

Physician members of DMCs are serving as consultants to the sponsor and, thus, are not covered by malpractice insurance that may be in force at the institutions where they practice. The same would apply to biostatisticians. All DMC members should be responsible for intentional negligence but indemnification has emerged as a standard way to protect DMC members against liabilities that arise in the trial for which they are not responsible. Confusion arises when a sponsor's legal department issues the same contract to DMC members as to investigators which usually would not have this

indemnification language. Writing the contract is the responsibility of the sponsor's legal department, but DeMets, Fleming, Rockhold et al. (2004) provide useful guidelines. Mutual indemnification, rather than just the sponsor's indemnifying DMC members, is often discussed. Most sponsors prefer just to indemnify the DMC members.

2.8 Sponsor DMC Relationship

The sponsor and DMC are both "officers" of the trial with the DMC being responsible for the stewardship. Neither is solely responsible for the trial. In order for the DMC to be able to fulfill its role, the tone must be set from the top. The sponsor is at the top; it is their trial, their money. The sponsor must respect the responsibility and authority of the DMC and must repeatedly acknowledge that they want the DMC to do a good job and help them in protecting patient safety and trial integrity. The sponsor should not intimidate the DMC by showing emotion as DMC recommendations are read. At the same time the DMC should not intimidate the sponsor with a "holier than thou" attitude or one of suspicion of sponsor motives. This relationship will evolve over time. It may not get off on exactly the right foot but it is the role of all players to make this important relationship work.

Sponsor attitude will contribute much to DMC morale, but the DMC has the advantage of being an ad hoc unit formed for a single purpose never to come together in this form again. Management experts have pointed out that teams formed in this way tend to have higher morale and are more efficient than teams of salaried employees with indefinite tenure (Blanchard, Carlos, and Randolph, 2001).

2.9 Interdisciplinary Training

While serving on a DMC, biostatisticians need to learn about the disease process, physicians need to learn about statistical methods used and both need to learn the details of the mode of action of the drug, pathways, cascades, kinetics, and so on. The DMC-sponsor relationship should be one of continuous on-the-job interdisciplinary training.

2.10 Conclusion

We now have learned about the within sponsor and outside sponsor members of the safety monitoring team and their roles. In the next chapter we will see how these players interact in meetings.

DMCounselor

Q2.1 I am a physician chairing a DMC for an Infant Pharma. The trial my DMC has been working on is about to come before an FDA advisory committee as a pivotal trial and I have been asked by the sponsor to present the efficacy results at the meeting. Given that I am a member of the DMC should I be doing this?

A You have hit upon an important point. As a DMC member you should not appear as an advocate for the product. That is for the sponsor and perhaps for other consultants. Explain to the sponsor that it is in their interest that they look for someone else. Perhaps you can help by recommending some appropriate consultants.

Q2.2 I am a biostatistician who was asked seven months ago about serving on a DMC for a phase III trial that was about to begin. I called the sponsor several times to reaffirm my interest. Each time I was told that DMC organization was to begin shortly. The sponsor finally sent me a contract yesterday, but I have also learned that the trial began two months ago and dose adjustments have already been made due to early adverse events. Should I join this DMC?

A Unfortunately this situation is not uncommon. This is a sponsor who may feel that the DMC implementation, although a requirement, is an annoyance and has made it of low priority. This may make it difficult for the DMC to have the appropriate stewardship. If you feel that you can educate the sponsor for improvement of their DMC operations, then it might be worth joining the DMC. If not, I would turn down their offer.

Q2.3 I am a project manager for a sponsor (Infant Pharma) trying to follow our SOPs in forming a DMC for our upcoming pivotal trial. NIH is a cosponsor of this trial and they are insisting that their standing 7-person advisory committee serve as DMC for this trial. There is nobody on this committee that we would have chosen for our DMC because the members have little or no drug development experience. We fear the committee members will continue to fulfill the functions expected of them by NIH but not provide the proactive monitoring required by the pharmaceutical industry. What should we do?

A It is good to have NIH as a cosponsor along with the prestige and credibility they can bring to this trial. Your company does not want to jeopardize these benefits with a dispute over the DMC. You must seek a win–win solution. How about proposing to NIH that your company will form a three-person DMC for the reasons that you have given and this DMC will make regular reports and share information with NIH's standing committee? Perhaps the NIH committee would consider making the chairperson of the DMC an ex-officio of the standing committee. It's worth a try.

Q2.4 I have been asked to serve on a DMC for a glaucoma product. The trial is a phase IIIB label extension trial. The product was approved several years ago and has been used by 100,000 patients postmarket. The sponsor wants to compare several schedules and drop certain arms at an interim analysis. However, I have learned that the trial has already begun and the sponsor will send us contracts and draft charter a few weeks before the first scheduled interim analysis. Sponsor claims there is no rush to do this paperwork because they already know a lot about the safety of this drug. The SOPs for forming DMCs are in place and are followed for pivotal trials, but the sponsor staff had made the decision that it is not necessary to follow the SOPs for a trial like this. Hearing this attitude has drastically reduced my enthusiasm for serving on this DMC. Am I being narrow-minded about this?

A No, your concern is appropriate. There may be some question of whether an independent DMC is needed for this trial. However, the sponsor has made the decision that they would like an independent committee to help them with certain decisions they will face in this trial. You and your fellow members should explain to the sponsor that given that they decided that an independent DMC would be employed on the trial, they should follow the same SOPs they have for pivotal trials. To have two standards is confusing for employees and their informal way of going about this can lower the standards for monitoring this trial which will result in an inferior trial from which nobody will benefit.

Q2.5 I am the biostatistician member of a DMC evaluating an experimental mood disorder drug sponsored by an Infant Pharma company. I have been working with the DAC biostatistician for six months now and I find that I am spending a lot of time teaching this person basic statistical methods. The DAC is a CRO and the biostatistician they have assigned is really more of a computer programmer than a biostatistician. I never expected to have to spend so much time on this and there is little return on the time that I spend because there are still many misunderstandings due to the lack of statistical knowledge. If I don't supervise this person, I don't know who would do so because the sponsor does not have a biostatistical staff.

A You should not be supervising the DAC biostatistician any more than you should be writing the statistical sections of the clinical trial report for the regulatory submission. It was the sponsor's responsibility to hire a qualified DAC and the DAC's responsibility to provide a qualified individual to serve as DAC statistician and to provide adequate training and supervision to this person. Your problem is that you need someone to talk to who understands your requirements. The Infant Pharma sponsor may not have a biostatistical staff but they must have biostatistical consultants. Similarly the DAC may have senior biostatisticians on staff or as consultants on retainer. You ought to tell the sponsor that you want to now communicate with someone like this who can, in turn, communicate requirements to and check the work of the DAC biostatistician.

Chapter 3

Meetings

Bullets to Remember

- The DMC charter is an important guideline to DMC operations in general and to meetings in particular.
- The orientation meeting is called by the sponsor. All subsequent meetings are called by the DMC in conjunction with the sponsor and in accord with the charter.
- The DMC should hold at least one face-to-face data review meeting each year. Remaining meetings during the year can be via conference call.
- Data review meetings have both open and closed sessions.
- Minutes of all meetings should clearly delineate sponsor, DMC, and DAC responsibilities for the next meeting.

3.1 DMC Charter

We have now learned about the purpose and formation of a DMC. Before turning to the structure of DMC meetings it is important to describe the *DMC Charter*. The charter is prepared by the sponsor and, as a result of SOPs, is very similar for all trials conducted by the sponsor except for trial-specific information. The charter will serve as a guide to meetings as well as the topics of the remaining chapters of this book. Table 3.1 lists the contents of a typical DMC Charter.

3.2 Types of Meetings

Many DMC meetings are a blend of several of the meeting types presented here. Some meetings are face-to-face and some are by telephone. The orientation meeting should be face-to-face. No face-to-face meetings should be held on sponsor's premises. Doing so can reduce the feeling of independence between the DMC and

TABLE 3.1: Items Typically Found in a DMC Charter

1. Formal name of clinical trial
2. Membership
3. Requirements for conflict of interest and confidentiality
4. DMC responsibilities
   a. Safety monitoring
   b. Efficacy data—interim analyses
   c. Publications
   d. Confidentiality
5. Chairperson responsibilities
6. Sponsor responsibilities and contact information
   a. Who is unmasked within sponsor, CRO?
   b. To whom does the DMC report?
7. Masking policy for DMC members.
8. Data Analysis Center responsibilities and contact information
9. Communication and data flow among DMC, sponsor and DAC
10. Software validation requirements, extent of data monitored at clinical sites before data review meetings.
11. DMC minutes and recommendations
12. DMC meetings
    a. Types of meetings
    b. Schedule of meetings
    c. Open and closed sessions
    d. Voting
    e. Masking policy
13. Procedures for recommending major changes to protocol
14. Resignation/termination of DMC member and replacement
15. Meeting minutes and retention
16. Safety analysis plan—templates of tables and listings to be reviewed during DMC meetings

the sponsor. There have been attempts to hold face-to-face meetings while DMC members are attending annual medical meetings such as the American Society for Clinical Oncology or the American Academy of Ophthalmology. In practice, many sponsors now avoid scheduling face-to-face DMC meetings at these annual congresses because of many time conflicts among members, some of which do not arise until the member arrives at the medical meeting.

3.2.1 Orientation or Organizational Meeting

This is the kickoff meeting where all of the players involved in DMC operations come together to review the DMC Charter and finalize a *Safety Monitoring Plan* (SMP). The DMC members will have reviewed the DMC Charter, protocol, and investigators' brochure prior to this meeting. It is highly advisable that this meeting be face-to-face.

3.2.2 Data Review

The *data review meetings* are those scheduled meetings where the sponsor will bring the DMC members up to date on trial operations, and the DMC members will have a closed meeting to review the safety data.

3.2.3 Ad Hoc

As the term implies, the ad hoc meetings are called for a specific purpose usually to address a safety concern.

In the next section we go into the nature of these meetings in detail.

3.3 Orientation Meeting

3.3.1 Chair for Orientation Meeting

Generally the sponsor will have appointed a member of the DMC to serve as chair prior to this meeting. Leaving the chairpersonship up to an election among members is not advisable. The sponsor must select a person with whom their staff can work and who can do the job expected. The chair and the sponsor representative will usually cochair the orientation meeting. All future meetings regardless of attendees or location will be chaired solely by the DMC chair. This is necessary to preserve the independence of the DMC.

3.3.2 Introduction of the Safety Monitoring Team

The sponsor representative will introduce all sponsor staff involved in the trial, as well as the DMC members and the DAC staff. The sponsor staff will usually consist of the study manager, others from clinical operations, and the internal safety monitoring committee, sometimes referred to as pharmacovigilence, firewall, or medical governance committee.

3.3.3 Appointment of DMC Secretary

The *DMC Secretary* can be a member of the DMC or the DAC statistician. This person will prepare the minutes of each meeting and circulate them for approval. Only DMC and DAC members will receive copies of the closed meeting minutes.

A plan for records retention should be made at this meeting. The records must be stored in a place to which sponsor staff associated with the trial will not have access until the trial is completed.

3.3.4 Presentation of DMC Charter

The cochairs will carefully go over the DMC Charter, accepting comments and revisions throughout the meeting. The typical sections of a Charter are presented in Table 3.1. We will be going through the items in the Charter throughout this book. The Charter represents the guidelines for operations and responsibilities of the DMC. It is as important as the protocol for the clinical trial.

3.3.5 Masking Policy

An important decision to be made at the orientation meeting is the policy of masking, which will be an item in the Charter. The Charter should indicate who at the sponsor can be unmasked. This will generally not be members of the sponsor's trial team but perhaps people in a pharmacovigilence group. It is important for the DMC members to know who can be unmasked at the sponsor because of important discussions of serious adverse events that may need to take place during the trial. The Charter will also indicate if the DMC will be unmasked to treatment assignment or partially unmasked where the members know the treatments only as A or B. There is much controversy on this matter, but it is agreed that DMCs may vote in closed session to be unmasked at any time. Those who advocate partially masked DMCs argue that there is no reason for anyone involved in the trial to be unmasked until absolutely necessary.

Many DMC members feel they can better understand the safety issues if they know which patients are being treated with the experimental group and which the control or placebo group. These people feel that if they were selected because they could be trusted with the stewardship of the trial, why they should not be trusted with knowing the identity of the treatment groups? In many trials DMC members can usually guess which treatment group is experimental after a few data review meetings just because of the pattern of adverse events. However, there is always the chance that they might guess wrong and, if so, this can upset decision making down the line.

3.3.6 Investigator Brochure

The sponsor representative will lead a summary of relevant sections of the *investigator brochure*. This document summarizes all data known about the compound under investigation. It will include safety and pharmacological data from animal studies as well as from earlier clinical trials. The investigator brochure will usually also include safety information on molecularly similar compounds. This document will suggest to the DMC members what safety issues to be on the alert for during the trial. At the conclusion of the investigator brochure discussion there should be agreement between sponsor staff and DMC members as to what adverse events are expected and what events that might be rare are nevertheless of interest.

In this era of rapid drug development sponsors will often begin phase III trials with fewer exploratory trials having been completed than has been the case in the past. This is definitely, but not exclusively, the case in Infant Pharma. This deficit in exploratory trials often prompts concern among DMC members about the dose and schedule being employed in the phase III trial. There is a definite impact of a sparse

investigator brochure on the mindset of DMC members and the safety monitoring procedures they will recommend. At one extreme, members may be more likely to terminate a trial due to toxicity in this case. At the other extreme they may operate more conservatively than otherwise especially if the drug is not in a class of drugs with similar molecular structure on which they can presume a safety profile.

3.3.7 Protocol

The sponsor representative will lead a discussion of the *protocol*. Here DMC members will be interested in eligibility requirements, the frequency of visits and evaluations, dosing and schedule, planned interim analyses, adverse event grading and coding conventions to be used, and so on. The DMC may recommend changes in eligibility, dose and schedule during the trial. Given the safety profile of the drug, the DMC members may recommend additional and/or more frequent diagnostic testing during the trial of such procedures as stress echocardiography, sophisticated scales such as the McGill Pain Index (Melzack and Torgerson, 1971) or the Hamilton Depression Scale (Hamilton, 1960). When multinational trials involve members from developing countries who have clinical trial experience in these countries, the DMC may want to inquire about how patient follow-up, patient compliance, SAE definitions, and so on will be implemented and monitored in these countries.

3.3.8 Informed Consent

Informed consent agreements are written by the institutions and approved by their IRBs. Sponsors usually provide a suggested wording providing their knowledge of expected adverse events. Most often DMC members will give their input on the contents of the sponsor's suggested wording at the orientation meeting. This wording will be revised as safety issues arise during the trial.

3.3.9 Data Flow

An important aspect of the orientation meeting is a discussion of data flow between the sponsor and the DAC, and the DAC and the DMC. The first item of discussion is the schedule for sponsor's sending DAC the data needed to fill the tables and graphs requested by the DMC. The second item would be the schedule of the DAC sending tables, listings, and graphs to the DMC prior to a meeting. Third item would be communication of SAEs to the committee. Table 3.2 presents a checklist for issues to be decided for this important communication. It should be noted that the orientation meeting is the beginning of a dialogue between the DMC and the sponsor on data flow requirements. It is not expected that all decisions will be finalized at this meeting.

This meeting will, at least, begin a dialogue between the DMC and DAC on statistical methods to be used in reports. Table 5.9 presents a list designed to help in this discussion. At the time the DAC begins implementation of the statistical methods into their computer programming, the members should issue a Statistical Analysis Plan (SAP) to the sponsor and DMC for approval and further comment.

TABLE 3.2: Issues for Discussion at Orientation Meeting on Serious Adverse Event Data Flow

1. What information will be in the report?
2. Who at the sponsor will be responsible for communicating SAEs to the DMC?
3. Which SAEs will be reported—all, only those possibly related and unexpected?
4. Will SAEs occurring on other trials not within the jurisdiction of this DMC be reported to the DMC? By whom? With what frequency?
5. Will the DMC review all deaths, only those related, only those occurring early in treatment? How will this review be handled?
6. Will this occur at any time or within 90 days of administration of study drug?
7. Will notification be as the requested SAEs occur or cumulative by week or month?
8. Will the SAE notifications first be sent to the DMC chair who will decide what further action is needed or will all DMC members receive the SAE reports at the same time?
9. Will communication of SAEs be via e-mail? If so, will this e-mail be password protected?
10. How often will these SAE reports be updated as new information arrives?

3.3.10 Useful Software

Many sponsors and DACs are now using Internet-based enterprise collaboration software to support DMC communication. This software allows DMC members to share information over the Internet. All documents needed by the DMC—investigator brochure, protocol, charter, SAE reports, tables, listings, and graphs—are posted at a secure Web site. Members are alerted by e-mail when new information has been deposited. This communication method saves paper and, because DMC members do not carry hard copies of documents, the chance of leaving a confidential document somewhere is reduced. Early experience shows that this approach is useful and is expected to become standard in a few years. When software of this type is to be used, a demonstration of its use at the orientation meeting is recommended. If a member prefers to receive documents by another means, such as paper reports sent via express mail carrier, this request should be granted.

3.3.11 Review of Integrated Summary of Safety

It must be decided at the beginning of the trial if the DMC will have responsibility for review of the *integrated summary of safety* (ISS). This is a compilation of safety data across trials in the clinical program. This document is created as part of the regulatory submission for marketing approval. DMC members will want to be sure that safety issues that they raised in their review of the confirmatory trial are not lost when the data are combined with smaller trials that may have had different patient eligibility requirements, doses and schedules, and so on. Weihrauch and Kubler (2002) and

Fairweather (1996) provide useful insight into the structure and use of the ISS. The sponsor may not want the DMC to review the ISS because it may be created some time after the DMC's final meeting and the trial team may not want the DMC to get on the critical path countdown to submission for fear of slowing down the process. Whether the DMC reviews the ISS may be a matter of discussion at the orientation meeting. The final decision on this responsibility must be made clear at the time of the forming of the DMC.

3.3.12 Policy on Review of Publications and Package Insert

If the DMC will have the responsibility for reviewing manuscripts on the trial that contain data reviewed by and decisions made by the DMC, this responsibility should be clarified at the meeting. A means for adjudicating disagreements on the contents of the manuscript should be formulated. Some DMCs request the right to submit their own manuscript or letter to the journal when there is disagreement. It is hoped that the latter will not have to be used, but some experienced DMC members feel that this is necessary in the event that sponsors underreport adverse events in publications. DMC review of manuscripts would be considered important to ensure that there is adequate reporting of incidence and severity of adverse events and reasons for discontinuation. Ioannidis and Lau (2002) provide useful guidelines for adequate reporting. Closely related to review of manuscripts would be the DMCs review of *package insert*. Although the DMC can comment on the sponsor's proposed package insert, the precise wording must be negotiated with the regulatory agencies, which have the final say.

3.3.13 Formats for Tables, Listings, and Graphs

The sponsor and DAC will present the mock tables, listings, and graphs that they are planning to prepare for DMC review. The DMC members will provide their own ideas. This is the beginning of a dialogue, but a schedule for finalization should be agreed upon at the meeting. Following chapters will describe useful data summaries.

3.3.14 Schedule First Data Review Meeting

The last item for an orientation meeting would usually be scheduling the first *data review meeting*. It is common to schedule a meeting after X patients have been enrolled or after X patients have been enrolled and completed Y cycles of treatment. However, it is important that a maximum time for first data review be scheduled in case enrollment is slow due to recruitment problems or more than expected patients have become ineligible. This is done so that early enrolled patients are not unduly at risk for lack of data review just because enrollment is slow. A better wording for first meeting schedule would be "after X patients have been enrolled and completed Y cycles of treatment or after Z months whichever occurs first."

Table 3.3 summarizes some of the agreements to be made at the DMC orientation meeting.

TABLE 3.3: Agreements to Be Made and/or Discussions Started at DMC Orientation Meeting

1. Designation of DMC chair if not previously decided
2. SAE data flow (Table 3.2)
3. Who will prepare minutes of the open and closed sessions and where will they be archived?
4. Modifications to the protocol, case report forms, informed consent
5. Suggestions on format of tables, listings, and graphs for data review meetings
6. DMC preferences for statistical methods and other details of the Statistical Analysis Plan
7. Will Internet software be used to update DMC members?
8. Masking policy for DMC members, sponsor staff
9. To whom does DMC report on routine basis when serious safety concerns emerge?
10. Plans for data quality control and software validation for data used for data review meetings
11. Will the DMC review publications, package insert, and integrated summary of safety?
12. Budget issues for DMC operations
13. Schedule for first data review meeting

3.4 Data Review Meetings

As a general rule there should be at least one face-to-face data review meeting each year. The others can often be handled by conference call, but face-to-face data review meetings may be more frequent depending on the complexity of the trial and developing issues. There are usually a total of 2 to 3 data review meetings each year but the frequency will depend on the specifics of a given clinical trial. At the time a data review meeting begins the sponsor will have a copy of tables, listings, and graphs of safety items with all treatments pooled while DMC members will have a copy with data presented by treatment group. The latter may be presented in coded form. Data review meetings consist of *open* and *closed* sessions.

3.4.1 Attendance

The entire Safety Monitoring Team will be present for the meeting—sponsor staff, DAC statistical representative and DMC members. The DMC chair will run the meeting and the DMC secretary will record minutes. Certain parts of the meeting will be considered open and others will be closed to all but DMC members and the DAC statistician.

3.4.2 Open Session

3.4.2.1 Study Progress

The chair will open the meeting, making sure that the minutes of the previous meeting have been accepted, and then call on the sponsor representative to make a report of study progress. This report will consist of enrollment progress, number of active investigator sites, adverse events of note for all treatment groups pooled, and so on. This will also be a time for the sponsor to call for protocol amendments, informed consent changes, investigator brochure updates to the attention of the DMC, and ask for advice and approval. The sponsor representative will also report on the progress of tasks initiated by the DMC at previous meetings.

3.4.2.2 Data Quality

It is understood that, at the time of a data review meeting, site monitoring will not have been performed to the extent for a completed trial, but the sponsor should report on the approximate percentage of adverse event data, laboratory values, and so on. A key part of this presentation would be the distribution of last date of contact for data to be reviewed in closed session. This will provide some idea of the currency of the data. If an SAE is under investigation and there have been, say, seven cases reported, but the last dates of contact range from 3 to 6 months previously, the DMC might want an updated analysis sooner than at their next scheduled meeting.

At this point DMC members will generally comment on any problems with the tables, listings, and graphs that they received in advance of the meeting. DMC members must be careful to speak only in general terms and not reveal trial data. This agenda item of the open session will close with the DAC statistician providing written certification that the computer programs used to generate the tables, listings, and graphs for the meeting have been validated according to industry standards. A list of standards employed by the DAC would be useful.

Glaser (2002) has provided some insight on quality criteria for statistical programming.

3.4.2.3 Update on Pending Action Items

The chair will then ask for a report of sponsor progress on issues brought up by the DMC at previous meetings. This may mean getting clarification on measurements from the central laboratory, writing a “Dear Investigator” letter, investigation of noncompliance, and so on.

3.4.2.4 Questions for the DMC

Many sponsors prefer to close the open session with specific questions that they want the DMC to take up in their closed session. The DMC may, and usually will, take up additional issues of their own choosing. Typical sponsor questions would be “Is there concern about the hypersensitivity reactions?,” “Is there any concern about mortality?,” “Can the trial continue without protocol modification?”

3.4.2.5 Sample Agenda for Open Session

Experience has found that agendas for open sessions that are merely lists of topics and do not clearly indicate responsibilities, the preparation needed, decisions to be made, and so on, have been found insufficient. An example of a useful format is displayed in Table 3.4.

3.4.3 Closed Session

At the closed session the DMC will meet together with the DAC statistician and go over the tables, listings, and graphs that were produced for the meeting. The DAC statistician will generally submit a data summary report to the DMC a short time before the meeting. It is best for DMC members to read this report after reviewing the meeting materials themselves. This is recommended to preserve the independence of the DMC and to ensure that members review material using the specific expertise for which they were appointed.

Members will review pending safety issues and seek evidence of new issues. The precise procedures to be followed will be described in the following chapters. The DMC has the right to hold an executive session in which all DAC members would have to leave the room for the duration of the executive session. At the conclusion of the closed meeting the chair will contact the sponsor representative to give a verbal overview of the meeting and indicate when the minutes of the open session will be available for review.

3.4.4 Scheduling of Next Meeting

The next meeting will be scheduled in accordance with the frequency indicated in the charter unless a meeting is needed sooner to respond to a safety issue or to get a briefing from a consultant with expertise outside of the realm of the DMC members. When the next regular data review meeting is scheduled the DAC representative will indicate a cutoff date for accumulated data that will be necessary for the DAC staff to prepare tables in time for the meeting.

3.4.5 Minutes

Two versions of the *meeting minutes* will be issued. One will cover the open session and be distributed to sponsor staff. This version should clearly describe sponsor, DMC and DAC responsibilities for the next meeting together with deadlines. The DMC version will cover both the open and closed sessions and include a list of pending safety issues that will need to be revisited until resolved.

TABLE 3.4: Sample Agenda for DMC Open Session

Date: February 26, 2008
Place: Chicago Airport Hotel, Chicago, IL
Attending: Sponsor: P.T., R.F.,* L.R., T.Z., J.C. DMC: C.L., R.A., V.L., S.R. CRO: A.Q.,* M.B. *By telephone
Meeting Objectives: Updates on study progress Updates on outstanding issues Sponsor informs DMC of specific questions they would like addressed
Preparation: Review minutes of last meeting—Group Create update reports, cutoff date 02/12/2008—T.Z. Review draft "Dear Investigator" letter–Group Research compliance situation among Latin American sites—A.Q. Create monitoring report, cutoff date 02/12/2008—A.Q. Create DAC report—M.B. Checklist of changes to reports agreed to at last meeting—M.B. Create list of questions for DMC—P.T.
Decisions to be made: Is enrollment satisfactory?—DMC Approval of "Dear Investigator" letter—DMC Has Latin American sites' compliance improved—DMC Changes to safety reports completed—DMC Monitoring, software validation sufficient—DMC Plans for next meeting—Group

(*Continued*)

TABLE 3.4: Sample Agenda for DMC Open Session (Continued)

Item	Time	Discussion Leader	Preparation	Decisions/Actions	By Whom	By Date
Attendance	10:00	C.L.				
Minutes of Last Meetings	10:05	P.T., C.L.	P.T.—Send minutes of last open meeting for review; C.L.—comment on approval and retention of closed meeting minutes	Approval of open meeting minutes	Group	02/26/2008
Trial Status Report	10:10	T.Z.	Obtain latest registration information by site, cutoff date 02/12/2008	Has enrollment improved? Additional sites needed?	Group	02/26/2008
Deaths and discontinuations due to study drug	10:30	T.Z.	Obtain pooled data, cutoff date 02/12/2008	Discuss in closed session	DMC	
Draft "Dear Investigator Letter"	10:45	J.C.	Write draft and circulate to DMC members no later than 02/19/2008	Discuss wording and send out	DMC	Mail final version by 03/03/2008
Compliance report Latin America	11:00	A.Q.	Obtain update on compliance in run-in meds at Latin American sites	Has there been improvement? Further action needed?	DMC	02/26/2008
Status of monitoring	11:15	A.Q.	Obtain monitoring levels of SAEs, death dates as of 02/12/2008	Is this adequate for closed session deliberations?	DMC	02/26/2008

Status of changes in reports requested by DMC at last meeting	11:30	M.B.	Use list of changes requested by DMC issued 10/31/2007	Have requests been satisfied?	DMC	02/26/2008
Status of software validation	11:30	M.B.	Obtain status of validation of software used for generation of tables at this meeting	Is this level adequate?	DMC	02/26/2008
Questions for DMC in closed session	11:35	P.T.	Meet with staff to draft list prior to 02/26/2008			
Closed Session	11:45–13:30	C.L.				
Open Session-Debriefing	13:45	C.L.				
Plans for next meeting	14:15	C.L., P.T.		Sponsor staff to make arrangements	P.T.	03/07/2008

3.5 Ad Hoc Meetings

The DMC may need to schedule *ad hoc meetings* to deal with emerging safety issues. These meetings will usually be run by conference call and will be closed session only. The agenda for the ad hoc meeting will be similar to that for the data review meeting. This meeting may be held without sponsor's knowledge, but if held with sponsor's knowledge, the sponsor would usually not know the agenda for the meeting.

3.6 Conclusion

This chapter has introduced the types and organization of DMC meetings. In our next chapter we will investigate the nature of the clinical data that the DMC reviews.

DMCounselor

Q3.1 I am a DMC chair. The sponsor just sent me and each of the DMC members an e-mail indicating that they are postponing our data review meeting scheduled for three weeks from now. They claim that the DAC has not had time to do the programming to prepare for the meeting. We don't want to postpone the meeting and we took a considerable amount of time to get our calendars in order for the scheduled meeting. Do we have to comply with the sponsor's change of date?

A The first issue is that the sponsor has no right to change the date of a DMC meeting. The DMC sets the dates and only the DMC can change the date. The sponsor should have contacted you, the chair, alone and tried to get your input on what could be done. The second issue is that the DAC should be more responsive to the requests of the DMC. There is usually more than enough time to do the programming between meetings. You may have no choice but to postpone your meeting but this would be a good opportunity to have a heart-to-heart talk to get the sponsor–DMC–DAC relationship on the right track.

Q3.2 I am a physician member of a DMC for a randomized active control trial in seasonal allergic rhinitis. The active control is a marketed drug for this indication. The tables we review show that there is a consistent 18% incidence of transient headache which we can confidently attribute to the experimental drug because we know that this does not occur at all on the active control. I am in favor of terminating the trial because I am confident that no physician is going to prescribe a drug with this side effect when the active control and other marketed drugs do not have this side effect. My fellow DMC members think such an extreme action would be out of line. What should we do?

A The issue you raise is a marketing issue and not a serious safety issue. The DMC is not charged with making marketing decisions, and it is not clear what the marketing potential will be once the efficacy data are known. Your concern is understandable but, for now, just concentrate on the serious adverse events. At the end of the trial I am sure the sponsor will value your input on marketing issues such as the one you raise.

Q3.3 I am an oncologist serving on a DMC for an oncology trial sponsored by an Infant Pharma company. At our open meetings the sponsor representatives take a lot of time asking the DMC members' advice on matters that do not concern patient safety. Many of the questions have to do with what licensing opportunities the company might pursue with Infant Pharma and if the company should be writing protocols to show the combined efficacy with a popular kinase inhibitor to increase licensing potential with that company. I do not feel comfortable or qualified to answer these questions but our chair and one other clinician member seem eager to discuss these matters. How do we get back on course?

A It appears that you have uncovered an important agenda item for your closed meeting. The DMC members are not contracted to provide such advice and it is thus out of scope. The chair is obviously not running the open session as he/she should be. If so, this type of questioning should not take place. At your closed session you should just remind the chair that these business topics are not appropriate DMC issues and try to build consensus to politely deflect these kinds of questions should they appear again.

Q3.4 I am the biostatistical member for a DMC working on an experimental drug used in emergency medicine. All clinics participating are emergency rooms. The other members of the committee are all ER physicians. We have had two telephone data review meetings so far and a face-to-face meeting is now scheduled for five months hence. The members are very competent in emergency medicine but as we look at the very nice safety tables produced by our DAC, our chair will interrupt with questions about waiver of consent and selection bias and the other members will begin giving their experiences with these issues. I have heard some of their experiences more than once even in the same meeting. As time passes beepers go off, we run out of time and I feel the meeting adjourns with not enough time spent reviewing the safety tables. How do I move this committee in the right direction?

A The good news is that the ER doctors on your committee definitely understand the concepts of doing clinical trials in emergency rooms. There are certain standards upon which informed consent can be waived due to the emergency nature of the encounter, and indeed, selection bias can occur because the most severe patients may be treated off-protocol with approved medicine immediately because there is not time to go through the enrollment and randomization process. However, there is no reason to be so obsessed with these issues that the DMC's safety responsibilities are compromised. It would be sufficient for the DMC chair to ask the sponsor team at each meeting if waiver of consent is being followed and if the sponsor's monitors have any evidence of patients who would be protocol-eligible being treated off-protocol. These are really sponsor responsibilities and not direct DMC responsibilities. You should diplomatically remind your chair of this and also use your DMC's Charter to show that these issues do not fit neatly into a DMC responsibility category. There is no problem with asking the sponsor for clarification of this matter should there be disagreement with your interpretation. However, there is no doubt that independent safety review is critical. It is possible that the physicians on your DMC might have looked at the safety data prior to the meeting, decided there was nothing serious to discuss and were thus comfortable curtailing the meetings. This, of course, is not acceptable and the minutes of your meetings should note that the safety data were not discussed in detail. You might want to request a "catch up" safety review by telephone rather than wait five months for the next meeting.

Chapter 4

Clinical Issues

Bullets to Remember

- DMCs use clinical data to separate signal from noise, that is, to differentiate adverse events associated with the study drug from those with other etiologies.
- A serious adverse event (SAE) is any untoward medical occurrence that, at any dose, results in death, is life-threatening, requires inpatient hospitalization or prolongation of existing hospitalization, results in persistent or significant disability/incapacity, or is a congenital anomaly/birth defect.
- An important sponsor policy is deciding who on the sponsor staff can be unmasked to treatment assignment in SAE reports.
- MedDRA (Medical Dictionary for Regulatory Activities) is the principal adverse event coding dictionary for pharmaceutical industry clinical trials.
- In multinational trials DMCs should understand cultural, political, and medical/surgical practice issues that may affect adverse event data.

4.1 Goals of Safety Analysis

The ultimate goal of safety analysis in clinical trials is to describe and evaluate patient risk for treatment emergent adverse events. To accomplish this, DMC members must separate adverse events that are part of the disease process, caused by preexisting or concurrent conditions or related to a concomitant medication from those that are related to study drug. In short, safety analysis seeks to separate signal from noise. To accomplish this goal DMC members will review tables and will occasionally use methods of statistical inference, but discussions of possibly serious treatment emergent adverse events will not be solely dependent on the result of a statistical hypothesis test. Statistical methods appropriate for DMC safety analyses will be covered in Chapter 5.

4.2 Definitions

Safety analysis entails continuous surveillance of many variables with many subclassifications in the effort to look for signals of risk. The need for common definitions and terminology in drug safety is a long-standing problem (Meyboom, Lindquist and Egberts, 2000; Aronson and Ferner, 2005), the World Health Organization (WHO) made an early attempt to standardize adverse event terminology (Edwards and Biriell, 1994) as did the Council for International Organizations of Medical Sciences (CIOMS; Venulet and Bankowski, 1998). ICH has published harmonized definitions for use in clinical trials for investigational drugs in their E2A guideline (International Conference on Harmonisation, 1994). The following definitions are paraphrased from E2A.

4.2.1 Adverse Event

An *adverse event* is any unfavorable and unintended sign (including an abnormal laboratory finding, for example), symptom, or disease temporally associated with the use of a drug, whether or not considered related to the drug. The term *treatment emergent* is often added as a modifier in order to remove manifestations of preexisting conditions from consideration.

4.2.2 Serious Adverse Event

A *serious adverse event* (SAE) is any untoward medical occurrence that, at any dose, results in death, is life-threatening, requires inpatient hospitalization or prolongation of existing hospitalization, results in persistent or significant disability/incapacity, or is a congenital anomaly/birth defect.

An important distinction is between a *severe* adverse event and an SAE. Severe refers to the intensity of the event but not necessarily to the seriousness. A patient my experience a severe headache, but it would not be considered serious by the above definition. Similarly a mild case of dehydration might cause hospitalization and thus be considered serious.

4.2.3 Serious Adverse Event Reporting Requirements

Regulatory agencies require expedited reporting of SAEs when they are unexpected. The latter would mean that there are no previous documented cases of this SAE for this drug either in the literature or in the investigator brochure. Investigators are asked to exercise judgment on whether the SAE was unrelated, possibly related, or related to study drug. Expedited reporting requirements vary between regulatory agencies but such reporting is often required for unexpected SAEs that occur within 28 days of study drug administration or after 28 days if the investigator deems them at least possibly related (U.S. FDA, 1997). IRBs and DMCs will receive expedited SAE reports. All other SAEs will be reported periodically to IRBs, the DMC and the regulatory agency.

4.3 Safety Data

4.3.1 Pharmacovigilence Group

Safety data will be processed at the sponsor in accordance with its SOPs and will be transferred to the DMC in accordance with the data flow plans in the DMC Charter. For Big Pharma and most of Middle Pharma, expedited serious adverse event data are processed through a separate pharmacovigilence group within the organization. This group is separate from the team working on the trial. The procedures of some sponsors allow or require this group to be unmasked to the treatment of the patient who experienced the event. Other sponsors insist on masking. In any event the team working on the trial should be masked. For the Infant Pharma companies, there is no clear pattern, but most choose for everyone in the company to be masked. This is done partly for scientific reasons but also to minimize the possibility of SAE-treatment information leaking to investors and the financial community.

4.3.2 Case Report Forms

All adverse events and laboratory values will be reported along with other clinical trial data on case report forms (CRFs). These forms will eventually be sent to the sponsor either electronically or by mail for processing for the final regulatory submission. The CRFs will record the description of the event, dates of onset and resolution, grade or severity of the event and the relatedness to study drug. The latter is open to investigator judgment. Definitions exist for determining grade and severity of an event, and these will be described below.

4.3.3 Adverse Event Dictionary

The adverse event reports must be coded in accordance with a dictionary. This is done by the sponsor's pharmacovigilence staff. The pharmaceutical standard for adverse event terminology is *MedDRA*—Medical Dictionary for Regulatory Activities. MedDRA was created in the 1990s as a joint venture between the International Federation of Pharmaceutical Manufacturers and Associations (IFPMA) and ICH. MedDRA is available through software on a subscription basis through the Maintenance Support and Services Organization (MSSO; MedDRA Maintenance Support and Services Organization, 2007). ICH provides periodic "Points to Consider" updates on use of MedDRA (International Conference on Harmonisation, 2007). The MedDRA dictionary (Brown, Wood, and Wood, 1999; Bousquet, Lagier, Lillo-Le Louet et al., 2005) provides a hierarchy of terms beginning with system organ class and progressing through preferred terms, high-level terms, and so on. For example, a system organ class might be "Blood and Lymphatic System Disorders" and preferred terms within this class might include *anemia*, *coagulopathy*, *eosinophilia*, *hypoprothrominaemia*, *thrombocytopenia*, and so on. The number of preferred terms generated within an organ class in an adverse events table will be referred to as the *granularity* of the table. Many of the adverse events enumerated for the marketed drugs in Appendix

Table A.1 represent combinations of several preferred terms. Sponsors usually follow a written internal policy for how combinations are to be made to define common adverse events such as headache, nausea, and dizziness. We will come back to granularity as a possible source of bias or inferential pitfall in Chapter 6.

In looking at Appendix Table A.1 we also note that heart blockage and bradycardia are listed as adverse events for cardiovascular drug metoprolol, angina pectoris for cardiovascular drug ramipril, and arthralgia for osteoporosis drugs risedronate and teriparatide. These were clearly adverse events observed in the trial but are also likely part of the disease process. Making these distinctions is an important role for DMCs.

4.3.4 Adverse Event Severity

It is essential that some system be used to attach a degree of severity to the adverse event. A coding system assigns an increasing AE intensity score called a *severity* or *grade.* Severity is generally coded as "none," "mild," "moderate," "severe," or "life threatening" and graded as 0, 1, 2, 3, or 4. "None" and "0" are the codes assigned to patients who did not experience the adverse event in question. Table 4.1 presents common definitions of severity some of which are in use by the General Clinical Research Center of the University of Washington (2007). Codes for grades are generally written for particular medical specialties using criteria based on data commonly collected in these fields. Examples are the U.S. National Cancer Institute's Common Terminology Criteria for Adverse Events (CTCAE; 2006), the U.S. National Institute of Allergy and Infectious Disease, Division of Microbiology and Infectious Diseases (2007), OMERACT 7 for rheumatology (Lassere, Johnson, Boers et al., 2005), the AEFI for vaccine trials (Bonhoeffer, Kohl, Chen et al., 2002).

For clinical trials in oncology the distinction between MedDRA and CTCAE is very important. CTCAE is a dictionary intended for use by oncologists in clinical trials that they may undertake in federal government sponsored trials. MedDRA is the dictionary of choice for safety reporting by FDA and other regulatory agencies.

TABLE 4.1: Severity Score

Severity	Definition
None	Adverse event not experienced
Mild	Transient, requires no special treatment or intervention, does not generally interfere with usual daily activities, includes transient laboratory test alterations
Moderate	Alleviated with simple therapeutic treatments, impacts usual daily activities, includes laboratory test alterations indicating injury but without long-term risk
Severe	Requires therapeutic intervention, interrupts usual daily activities
Life threatening	Requires significant therapeutic intervention and patient is at *immediate* risk of death.

Source: University of Washington (2007), *Adverse events: definition and grading*, http//www.crc.washington.edu/DataSafetyMonitoringPlans/Adverse.aspx.

Some pharmaceutical industry sponsors of oncology trials specify CTCAE grades to be used with MedDRA terms. Oncologists serving on DMCs may need some orientation to MedDRA if they are used to thinking in CTCAE terms. Pharmadhoc (2007) has provided useful MedDRA—CTCAE mapping tools.

4.3.5 SAE Narratives

The coding conventions described above have been created for use in generating statistical tabulations of adverse events in a manner that we will describe in the next chapter. When deaths and serious adverse events occur physician members of DMCs want to see narrative descriptions of the events. These descriptions are written by pharmacovigilence staff, masked to treatment assignment, after reviewing all information available including interviews with the investigator whose patient experienced the event. The narrative will describe the event by the MedDRA-preferred term giving patient age, gender, date of onset, investigator opinion on relatedness, medical history, relevant baseline information, comorbidities, concomitant medicines, and so on. The content of these forms is often discussed by DMC members during closed session. These forms are also sent to regulatory agencies when the events qualify for expedited review. The narrative form is updated as new information on the event arrives. Some sponsors have their own narrative report format but the standard format used by most companies is that developed by CIOMS (2007).

4.3.6 Titration to Dose

In some protocols patients are titrated to dose—that is, the investigator progressively increases the dose until a dose that is "optimal" for that patient (using efficacy response and tolerability criteria) is reached. It will be important for the DMC to ascertain that the protocol for the trial provides a detailed description of how titration is performed in sequential visits, that the dose for each appears on the case report form and on SAE narratives, and that investigators are complying with the titration protocol.

4.4 Deaths

Deaths are an important consideration in DMC review. The DMC will attempt to determine if deaths are due to the disease process or to the drug and, usually, not rely solely on a comparison of treatment groups especially when an active control group is used for a life threatening disease, such as cancer. Deaths might stand out more in an allergy trial than in an oncology trial. Even in the latter the early deaths, those occurring in the first cycle of therapy, might be suspicious. Oncology DMCs use the term "death as a first event" or "death less than 30 days post–treatment start" to describe this phenomenon and seek reports of these early deaths. Most DMCs will not limit their review to deaths that the investigator or sponsor pharmacovigilence unit classify as drug related. This would violate their mission of stewardship. Johann-Liang, James, Behr et al. (2005) discuss this issue in relation to deaths on

HIV clinical trials. Quite a bit of judgment and second guessing is involved in these deliberations. We will return to this topic in later chapters.

4.5 Impact of Multinational Trials

The past decade has seen an accelerating trend toward performing clinical trials globally especially in India, China, and Eastern Europe. Pivotal trials today can include investigators in Moscow or Mumbai as frequently as Memphis or Montreal. Sponsors find lower costs and higher level of trial participation in these regions than in North America. These countries have large numbers of untreated patients eager to enter trials because this is often the best route to medical care. Many investigators in these countries are Western trained, they have clinics built specifically to conduct clinical trials, they adhere to ICH guidelines, and sponsors can hire or contract with physicians (rather than nonphysicians as is the case in North America) to act as clinical research associates to monitor protocol and regulatory adherence at sites (Platonov, 2003; Kahn, 2006). Hence, there is motivation among sponsors to include sites in these developing countries. DMCs should be aware of varying data quality due to the lack of long experience in pharmaceutical industry trial participation and other cultural differences discussed below.

The following sections review some global issues that impact the work of a DMC.

4.5.1 Cultural Issues

Geographic patterns of genetic variation have been known to affect adverse drug reactions (Wilson, Weale, Smith et al., 2001). Diets followed by some countries could interact with experimental drugs to give the appearance of drug-related AEs. Different cultures have different propensities to self-report the AEs, and these personality issues are further impacted when caregivers are the chief reporters of AEs to investigators as is often the case in Alzheimer's disease and Parkinson's disease.

4.5.2 Political Issues

Some state-run European health systems present physicians with financial incentives to hospitalize patients (Haluska and Aamdal, 2007). The hospitalization automatically qualifies an adverse event as an SAE. Thus we may see more SAEs reported from these countries. DMC members should not discount SAEs being reported from any country, but this is something to take into consideration. Some European regulatory agencies require sponsors to unmask whenever an SAE occurs (Stump, 2007). Although the FDA does not support this practice, sponsors must adhere if their trials come under these jurisdictions. Other political issues sometimes encountered would include patient attitudes due to suspicion of the capitalist system and difficulty with the concept of informed consent.

4.5.3 Medical/Surgical Practices Issues

Although Western medicine methods prevail in these new clinical trials markets, national health system formularies occasionally differ from North America in the nature of supportive care—use of anti-infective drugs, antiplatelet drugs, and so on, which can affect the level of adverse events. Similarly, funding of surgical techniques may differ and this can be a factor when study drugs are used postsurgery. For active control trials, the active control drug may vary over countries, due to international differentials in approved drugs or drugs supported by national health systems. Also, ethical requirements usually indicate that for certain diseases the control group must be standard of care. However, standard of care varies between regions and DMC members may be reviewing data from trials that have varying control groups depending on investigator location.

Table 4.2 reviews the multinational issues and provides suggestions for DMC investigation and response. The table mentions the possible use of logistic regression analysis. This method will be discussed in the next chapter.

TABLE 4.2: Issues in Multinational Trials

Issue	Possible DMC Response
Cultural	
Geographic genetic variation	Ask sponsor for literature search of what is known about genetic variation related to this disease and treatment. If this is a factor, a stratified analysis or covariate-correction through logistic regression might be requested.
Differential diets among countries	Ask sponsor for literature search of how diet might affect AE levels. If there appears to be an effect, a stratified analysis or covariate-correction through logistic regression might be requested.
Propensity to self-report AEs	Compare incidence of AEs that are self-reported among countries. If there appears to be an effect, suggest guidance to the investigators and request stratified analyses.
Political	
Financial incentives to hospitalize patients	If there appears to be a higher SAE level for an AE type in some countries than others, investigate if these are countries with financial incentives to hospitalize. If so, review the CIOMS forms to see which SAEs can be classified as AEs for DMC purposes only. If there are some, perform a DMC analysis. This analysis will not be considered the official drug application analysis.

(*Continued*)

TABLE 4.2: Issues in Multinational Trials (Continued)

Issue	Possible DMC Response
Unmasking required when SAE occurs	This should have little effect because the SAEs probably de facto unmask anyway.
Suspicion of the capitalist system	Investigator staff must reassure potential patients of the benefits of the clinical trial.
Difficulty with the concept of informed consent	Investigator staff members need to spend time with the patients and their families to make sure the informed consent form is well understood. Patients concerned about the potential adverse events should not be enrolled.
Medical/Surgical Practice	
Use of supportive care	Through a "Dear Investigator" letter, try to harmonize supportive care. When differences exist, consider their impact. Stratified analysis or covariate-correction through logistic regression might be needed.
Different surgical techniques used	This may be difficult to harmonize because of differing training, skills, and equipment among countries. Stratified analysis or covariate adjustment through logistic regression may be required.
Active control may vary over countries	Perform an analysis to see if odds ratios vary between active controls for the same AE type. If so, correction can be made through covariate adjustment in logistic regression.
Control group must be standard of care but standard of care varies among countries	Perform an analysis to see if odds ratios for the same AE type vary among countries that vary for standard of care. Correction can be made by covariate analysis in logistic regression.

4.6 Conclusion

We have seen that there is considerable subjectivity in reporting and classifying adverse events and these factors are compounded in multinational trials. These factors will affect DMCs' mission in separating signal from noise and in making safety decisions. There is no one perfect way to handle this situation, but we will learn more about current practice in the next chapter on statistical methods.

DMCounselor

Q4.1 I am an infectious disease specialist serving on a DMC for an experimental infectious disease drug. The committee agreed that we need an electrocardiologist consultant to help us interpret some the severity of some cardiovascular SAEs. The consultant selected by the sponsor wants to hold a half-day in-service training with slides he uses to train his residents. We don't think this is necessary.

A It sounds like the consultant believes in one-size-fits-all teaching. Although this electrocardiologist is paid by the sponsor he/she reports to the DMC, and thus, the DMC must decide how his/her time is best spent. It is the DMC's responsibility to write out the objectives and scope of the consultant's engagement. Does the DMC really want to learn electrocardiology or do they want the electrocardiologist to look at the data on some adverse events and advise on the seriousness? It might be good for the consultant to provide some guidelines for the DMC members to judge future SAEs of this type, but it might be more prudent to have the consultant return to review more cases as they develop. The key here is that the DMC must take the lead and be in charge of this and not follow what the sponsor and their favorite electrocardiologist think is right. If your DMC cannot work out an acceptable agreement with this electrocardiologist, you might want to seek another consultant.

Q4.2 I am a physician and chair of a DMC for a rheumatology indication. We have asked the sponsor for more information from some investigators on SAEs that we have been reviewing. The information has not arrived. How long should we wait before we give up?

A You are observing that the DMCs operate in an imperfect world. It is the sponsor's responsibility to get this information for your committee and its representatives should be reporting to you on their progress during open sessions of your meetings. However, the sponsor too works in an imperfect world. They may not be able to get further information because such information may not exist or the investigator is just too busy to send it. If the information does not come between meetings of your DMC, you must assume that it will never come and do the best you can with the information you have while continuing to keep the request open with the sponsor until they satisfy the DMC that they have tried everything and there will be no more data.

Q4.3 The sponsor for our DMC has opted not to use a central lab for the Eastern European sites participating in this trial due to logistical problems. The North American sites are using a central lab and those in India are using a Bangalore-based branch of the North American lab. The serum chemistries coming from the Eastern European sites have a much higher incidence of abnormal values than those from the other sites. These sites are not declaring the abnormal values as AEs which is troublesome enough but we are spending a lot of time talking about the lab values from these countries and, given that their origin is different from the other investigator site, we are probably wasting a lot of time.

A Welcome to the world of multinational trials. The eastern European sites all have separate labs at their institutions. They probably have different equipment and different ways of determining the laboratory range of normal. Your DMC should review the reported AEs and SAEs of these sites carefully and ask the sponsor to find out if the patients with the abnormal values have associated symptoms that would be expected for these lab values such as neutropenia, thrombocytopenia, or liver function. You can press the sponsor for this type of clinical information because it is likely to exist.

Q4.4 I am the chair of a DMC working on a treatment for inflammatory bowel disease. The trial began five months ago. To increase enrollment the sponsor has added a Canadian site and is considering some eastern Europe sites. The Canadian site has told the sponsor that they will not participate unless a Canadian physician is added to the DMC. My fellow DMC members and I see no need for another member, and we have concern that we will soon be asked to add an eastern Europe representative. Isn't this getting out of hand?

A The DMC must be part of any decision to add a member. However, it sounds as if the sponsor has been honest with you on the reason for adding the Canadian member. Rather than resist the addition of a Canadian member, I suggest you consider the advantages of having regional representation in multinational trials. In addition to clinical knowledge, this person brings an understanding of the practice of medicine in Canada, payment systems, regional nutritional trends, and so on, all of which could influence adverse event frequency. The DMC has a right to assess whether the candidate can provide this type of information. Eastern Europe is much different from North America, and your DMC may benefit from a representative from this region if sites are to be established there. Of course this can get out of hand if Latin American and African sites are also to be added. At that point the sponsor might agree to add the additional members only after six-month accrual quotas are met. If these quotas are not met, there would be less need for the additional regional members.

Chapter 5

Statistical Issues

Bullets to Remember

- A statistical calculation such as p-value should not be the sole criterion for eliminating AE types from further discussion; clinical significance should trump statistical significance.
- Safety analysis should be performed on the intent-to-treat population.
- It is important to take drug exposure into account while assessing treatment differences in AE incidence.
- The Kaplan–Meier graph is a convenient way to view AE incidence over time.
- The statistical basis of frequentist statistical methods lies in repeated sampling, that is, variation introduced by repeating the clinical trial an infinite number of times.
- Confidence intervals, likelihood support intervals and Bayesian credible intervals all provide a range of plausible values of a parameter although with different interpretations.
- Odds ratios and Poisson rate ratios indicate relative risk of AE occurrence between two treatment groups.
- Calculation of p-values for many AE types exposes inference to multiplicity whereby some AE types will have statistically significant p-values due to chance alone.
- The False Discovery Rate (FDR) is a means of correcting for multiplicity.

5.1 Goals of Statistical Analysis

The objective of this chapter is to review some common methods of statistical analysis that would be of use to DMCs and to demonstrate their interpretation in the context of DMC operations. Some novel methods that are of particular use to DMCs are presented. This chapter is not meant to be a statistical textbook; thus we will not reproduce formulas that are easily available elsewhere and where calculations are commonly made by readily available statistical software. Details are provided in a

nonrigorous manner for certain techniques that are not as well known and may be especially of interest to the biostatistician reader.

The ultimate goal of safety analysis in clinical trials is to describe and evaluate patient risk for treatment-emergent adverse events. To accomplish this, DMC members must separate adverse events that are part of the disease process, caused by preexisting or concurrent conditions, or related to a concomitant medication from those that are related to study drug. In short, safety analysis seeks to separate signal from noise. To accomplish this goal, DMC members will review tables and will occasionally use methods of statistical inference. Unlike efficacy analysis of primary endpoints, inference will not hinge on the statistical computation of a single endpoint. Safety analysis entails continuous surveillance of many variables with many subclassifications. There will be numerous important discussions of possibly serious treatment-emergent adverse events regardless of the result of a statistical hypothesis test. The role of the DMC biostatistician will be to remind physician members of the weight of evidence taking uncertainty into account. These are among the concepts we will discuss in this chapter.

5.2 Useful Data Displays

This section will describe certain types of tables, listings, and graphs that have been useful in DMC data review meetings. Specialized tables may be needed depending on the indication. For any DMC the data displays reviewed should be considered a work in progress—consideration of modification should be continuous in order to properly support emerging issues, and certain displays may be discontinued because they are deemed no longer necessary.

The DAC will prepare two versions of each table, listing, and graph. One version, which will be for use by the sponsor, will have all treatments pooled. The other, for DMC use, will be presented by treatment coded as A, B, C; Blue, Green, Yellow; or Tulips, Roses, Orchids. This presentation is known as *partially masked.* If the DMC is to be unmasked, a partially masked table will still be presented with the independent statistician decoding during the closed session. Partial masking is necessary in case the tables are inadvertently left in a hotel or airport lobby. More will be said about masking policy in Chapter 7. For partially masked tables, the treatment codes should be consistently matched with actual treatments at each meeting of the DMC (i.e., A, B, C, etc., should represent the same treatments at each meeting).

It is assumed that all data displays will be performed for the *intent-to-treat* population of patients, meaning all patients randomized regardless of how much treatment they received during the trial. In some cases it may make sense for analysis to also concentrate on the *adherers only* subset, which includes those patients who received treatment according to the protocol or some other criterion (Piantadosi, 2005).

Table 5.1 presents a list of useful data displays. Note that the first two displays listed deal with patient enrollment. These are for open session discussion and would not generally be produced on a treatment group basis. An exception would be if there

TABLE 5.1: Useful Data Displays

1. Patient enrollment by center (not usually produced by treatment group)
2. Cumulative patient enrollment by month (not usually produced by treatment group)
3. Cumulative distribution of patient exposure
4. Table of reason for discontinuation
5. Data currency table—data management cutoff date for this meeting, most recent death, and SAE by center
6. Treatment emergent adverse events by body system, possibly subclassified by
   a. Grade
   b. Relatedness to study drug
   c. Event types within body system
7. Same as 5 for serious adverse events
8. Laboratory values of interest. List by patient and flag those outside of normal range.

Note: Sponsor version for both treatments pooled, DMC version by coded treatment group.

were a need to consider whether randomization was flawed (unbalanced by treatment group). The following is a description of each display.

5.2.1 Enrollment by Center

A listing of enrollment by center including a calculation of enrollment/month of center participation is useful. For trials that have a run-in screening phase where patients qualify for randomization (see, for example Faught, Sachdeo, Remler et al., 1993, in epilepsy and Thijs, Celis, Kiowski et al., 1995, in hypertension), this should be calculated for both phases. DMC members may comment on nonperforming sites but also, in multinational trials, get an idea of from what part of the world the patients are coming.

5.2.2 Graph of Cumulative Patient Enrollment by Month

This graph (example Figure 5.1) will show the total patient enrollment by months since the start of the trial. Some sponsors prefer to add the monthly estimated patient enrollment (usually a straight line) to the graph. This will enable DMC members to review enrollment progress over time. For trials having screening or run-in phases, a line can be added for the qualifying phase. The estimated patient enrollment will not always be a straight line. The estimated enrollment will take different shapes if patients must experience a relapse on a previous trial in order to qualify for enrollment, in infectious disease trials where there is seasonal variation in incidence of the disease under investigation or where the sponsor plans to phase in sites over time.

5.2.3 Graph of Cumulative Patient Exposure to Study Drug

As DMC members discuss adverse events in closed session, it is important for them to know how much exposure patients have had to study drug. If there is concern for cardiac toxicity, for example, inspection of this graph may show that, because only

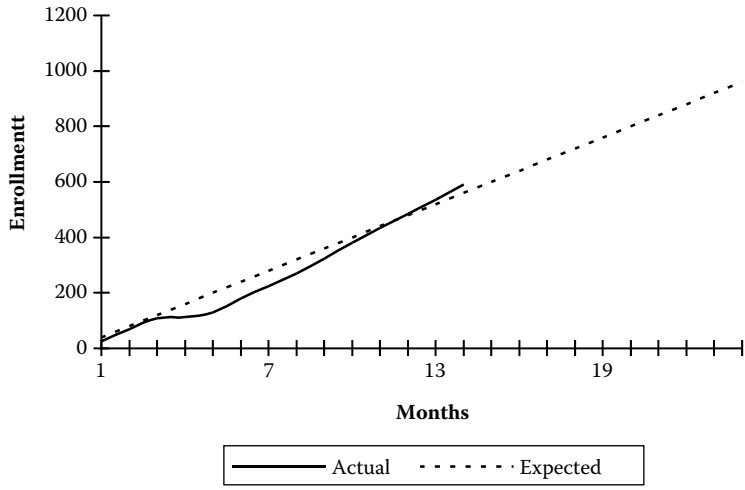


Figure 5.1: Cumulative patient enrollment by months since the start of the trial.

20% of patients have received three cycles of the drug, it is too early to dismiss cardiac events as a concern. If later in the trial, after 70% of patients have received three cycles of the drug and only 2% of patients reported cardiac toxicity, DMC members may begin to feel confident that cardiac toxicity will not be an issue. An example of a cumulative patient exposure graph is shown in Figure 5.2.

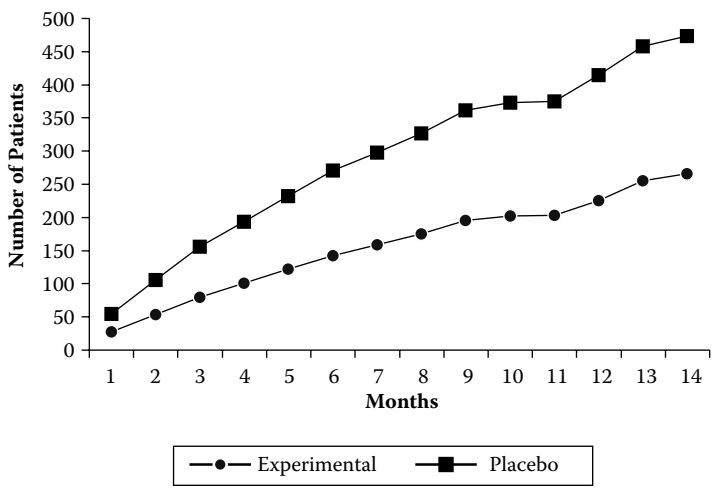


Figure 5.2: Cumulative patient exposure by treatment group and month since treatment start.

5.2.4 Treatment Emergent Adverse Events

5.2.4.1 Classification

There are many ways to classify and analyze adverse events. The art of classification has been discussed in Chapter 4 and classification as a source of bias will be covered in Chapter 6. For now we will concentrate on body system, subclassification within body system, relatedness to study drug, and grade or severity. An investigator's conclusion about relatedness to study drug is a regulatory requirement. In order to be conservative many DMCs ignore this variable in analysis especially in open label trials. The first line of any AE tabulation should be the summation of AEs for all body systems combined. This will give DMC members a sense of how many patients are experiencing AEs of any kind.

For each treatment group the number of patients and the total exposure for the group should be presented. Exposure will often be patient years of follow-up, but in some trials number of cycles or number of injections might be considered more relevant to safety.

Now the table proceeds to list adverse event types by body system and preferred term within body system. At the start of a trial the degree of granularity of preferred terms within body system should represent the DAC's best guess of an appropriate level. The DMC can request revisions in the preferred terms after some experience accrues. The art, science, and potential biases of granularity selection will be discussed further in Chapter 6. It is by no means an unimportant issue. For each AE type, when the patient has experienced several episodes of the same AE, the patient should generally be presented only once classified as the most severe episode. Most tables will have columns indicating the frequency of AEs by grade or severity. DMC members may prefer to consider only incidence of AEs classified as grade 3 or greater or moderate or worse. This would be especially true in life-threatening diseases.

Several DACs are now beginning to place MedDRA code numbers in parentheses after the text definition of each preferred term. This is a good practice. As experience with the MedDRA hierarchy begins to pervade the physician–biostatistician community, these numbers will be useful to DMC members.

There are two types of calculations that can be made from treatment-emergent adverse events from tables such as this one.

5.2.4.2 An Example

We will illustrate several analytic techniques with the example of cardiovascular events associated with rofecoxib as presented in Bresalier, Sandler, Quan et al. (2005). This trial was commonly known as the APPROVe trial (Adenomatous Polyp Prevention on Vioxx). Some controversy exists for this trial and several papers were published after the original providing correction to some data. We use the original data because it is being presented solely for illustration and not to present a point of view regarding the safety of rofecoxib.

Table 5.2 presents the data on adjudicated (confirmed) thrombotic adverse events from APPROVe that we will use in the following sections. The primary objective of

TABLE 5.2: Adjudicated Thrombotic Adverse Events in the APPROVe Trial

	Rofecoxib	Placebo
n	1287	1299
Patient years	3059	3327
No. of events	46	26
Incidence (%)	3.57	2.00
Rate/100 patient years	1.50	0.78

the trial was to determine the effect of three years' treatment with rofecoxib on the risk of recurrent neoplastic polyps.

The trial randomized 2,586 patients with a history of colorectal adenomous polyps to either rofecoxib or placebo—1287 to rofecoxib and 1299 to placebo. Rofecoxib patients contributed 3059 patient years and placebo patients 3327.

5.2.4.3 Incidence Calculation

The incidence calculation is merely the percent of patients in each treatment group who had experienced an adverse event—for rofecoxib 3.57% (46/1287) and placebo 2.00% (26/1299). We call the proportion of patients who experience the AE *proportional incidence*. For rofecoxib proportional incidence is 0.0357 (46/1287) and for placebo (0.0200). It is important to note that in all safety calculations each patient is counted only once. If there were repeat episodes of the same AE type for a patient, the patient is counted in the numerator in an incidence calculation only once.

5.2.4.4 Incidence and Exposure Time Calculation

The overall incidence by treatment group calculation would be appropriate if patients in each of the treatment groups had equal exposure to study drug. This will not be the case when there is a differential in dropout rates between treatment groups. The difference might be due to toxicity, so it is important to take exposure into account.

Exposure time and incidence can be combined in either of two ways. The first is to calculate the incidence per unit time such as number of events per 100 patient years. The second is a graphical method discussed in the next section. The rate per unit time calculation is made by dividing the number events by the total number of patient years on the trial and multiplying by 100. In our example the rate per 100 patient years for rofecoxib is 1.50/100 patient years $(46 \times 100)/3059$ and for placebo it is 0.78/100 patient years $(26 \times 100)/3327$.

The question arises whether events should be counted after the patient has experienced the first occurrence of a certain AE type and whether repeat episodes of this event should be counted. The most common convention is to count patient exposure as long as the patient remains on the protocol. To do otherwise would present different numbers of patient years for each AE type and there appears to be no benefit from doing this. At some point the DMC may want to concentrate on a certain AE type and then it may be appropriate to ask the DAC to compute rates using exposure only

up to the first occurrence of that AE type. However, as we will see below, the Poisson distribution (Rosner, 2006) is the reference probability distribution used for statistical inference of the rate per 100 patient years, and this distribution would, theoretically, require all occurrences and all exposure time to be counted in the numerator and denominator, respectively. This convention is not generally followed in pharmaceutical trials, but it is standard in adverse event analysis of clinical trials of mechanical heart valves (Grunkenmeier, Johnson, and Naftel, 1994).

First occurrence and exposure are more accurately computed in the following section.

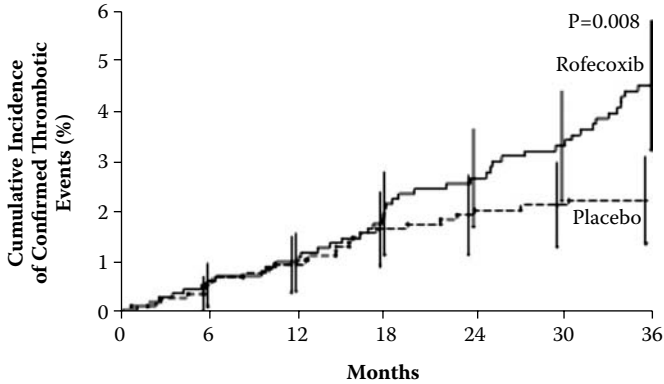
5.2.4.5 Kaplan–Meier Time to First Occurrence

Kaplan–Meier time-to-event curves, which provide a graphical view of the occurrence of events over time, take the number of patients at risk at each time point into account by dropping patients who discontinue or experience the event along the way. For patients who do not experience the event, their "time to event" is just their time on study (their total time) and this observation is called *incomplete* or *censored.* Those patients who do experience the event have their time to event recorded. Their observation is called *complete* or *uncensored*. The Kaplan–Meier methodology, formally known as the *product limit method,* is often applied to efficacy analysis but can easily be applied to safety data. In oncology trials the method is frequently applied to survival analysis (time to death). However, a patient can die only once but can experience the same type of adverse event more than once. We use the Kaplan–Meier analysis for adverse event data to describe the *time to first occurrence* of the event. The methodology for calculating a Kaplan–Meier curve is beyond the scope of this book, but details can be found in numerous references and software is readily available (Kaplan and Meier, 1958; Piantadosi, 2005; SAS Institute, 2008; Cleves, Gould, Guttierrez et al., 2008). Figure 5.3 displays the Kaplan–Meier curve for time to adjudicated thrombotic adverse events in APPROVe. The vertical axis indicates the cumulative number of thrombotic events, and the horizontal axis indicates time. This graph suggests that the thrombotic AE experience is about equal in both groups but that the curves begin to diverge after 18 months with more events accumulating for rofecoxib than for placebo.

Kaplan–Meier time-to-event curves would not normally be routinely generated for all adverse event types. That would create voluminous output. These curves would be requested by the DMC for AE types of particular interest and usually to study the development of these events over time.

5.2.4.6 Incidence at a Time Point after Treatment Start—Landmark Estimate

It is not uncommon for DMCs to request a *landmark estimate* of incidence (sometimes referred to as a point *estimate)* from a Kaplan–Meier life table. This estimate consists of specifying the Kaplan–Meier estimate of AE incidence at a particular time point such as 12 months after treatment start. This is an estimate of cumulative incidence from treatment start up to and including 12 months post-treatment start. The estimates and their standard errors are found on listings of the Kaplan–Meier *life table* (Kaplan and Meier, 1958; Rosner, 2006; SAS Institute, 2008; Cleves, Gould, Guttierrez et al., 2008). They can also be read from a Kaplan–Meier graph. From Figure 5.3 we see that



No. at Risk							
Rofecoxib	1287	1129	1057	989	938	896	727
Placebo	1299	1195	1156	1079	1042	1001	835

Figure 5.3: Kaplan–Meier estimates of the cumulative incidence of adjudicated thrombotic events (vertical line indicates 95% confidence intervals). *From:* Bresalier, R.S., Sandler, R.S., Quan, H. et al. (2005) Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial, *New England Journal of Medicine*, **352**, 1092–1102.

in the APPROVe trial both treatments have a 12-month incidence of serious thrombotic events of about 1%. This type of analysis is useful to show that certain AE types occur mostly in the first, say, six months after treatment start and then level off. It is important that landmark estimates of incidence be made using Kaplan–Meier methods. Computing the raw percentage of patients who experienced an adverse event within the first, say, 12 months of treatment has an obvious numerator, but the denominator is difficult to compute due to patients who are censored before 12 months.

5.2.4.7 Other Ways of Looking at Incidence

After viewing some SAEs, it is possible that DMC members may feel that the SAEs are related to a dosing of the drug. In these cases DMCs might request an additional table of SAEs that occurred within a time window (such as 30 days postinjection).

We will use the term *covariate* to mean a variable that might influence AE incidence. In most cases raw incidence rates will be sufficient for DMC purposes. However, in some cases AE incidence may vary by geographic region due to multinational issues discussed in Chapter 4. As one example, the control group in a clinical trial might be standard of care but that standard varies by geographic region. The statistical method of *logistic regression analysis* (Hosmer and Lemeshow, 2000) can be helpful in this case. This method allows for assessing the effect of covariates on incidence. In our example standard of care and treatment might be included as covariates in a logistic regression model. A statistical test of significance can be made to see if standard of care has an effect on AE incidence over and above treatment. If the effect is statistically

significant, standard-of-care-adjusted rates can be computed and compared between treatments.

5.2.5 Laboratory Data

DMCs are often provided with listings of laboratory data—serum chemistry and urine analysis. The specific use of these data by DMCs depends on the disease and intervention. The clinical laboratory data may be of use to find the effects of treatment on analytes that may be related to renal function, heart function, anemia, and so on. The sponsor should establish laboratory ranges of normal and clinically significant changes for each blood and urine analyte at the commencement of the trial. These parameters may differ by patient age, by gender and among laboratories if more than a single central laboratory is used. Changes in analytes from baseline or from visit to visit that are deemed clinically significant should be reported as either AEs or SAEs. Reference to a patient-by-patient listing of laboratory values and changes from baseline by DMC members will show if investigators are complying with this requirement and may turn up safety signals that are not apparent from AE and SAE reports. As the trial progresses the DMC may wish to concentrate on particular analytes of interest and ask the DAC for listings and graphs that illustrate the dynamics of these measures over time.

5.3 Analysis Methods—Frequentist

5.3.1 What Is Frequentist Analysis?

Frequentist analysis is the statistical methodology that is commonly taught in statistics courses, described in textbooks, and used extensively. This application of statistics depends on repeated sampling for its measure of uncertainty and inferential basis. We will illustrate how repeated sampling plays into the frequentist statistical methods we use as we introduce each concept. We use the term *frequentist* here to distinguish these well-known methods from likelihood and Bayesian methods which do not depend on repeated sampling for inference. We will discuss these methods later in this chapter.

5.3.2 Hypothesis Tests

DMCs will often be presented p-values for hypothesis tests on AE incidence. The p-value is also known as the *attained significance level.* In statistical terms the null hypothesis always states the negative result, so the null hypothesis for comparing AE incidence between two groups would usually be that there is no difference in AE frequency between experimental treatment and control in the target patient population. The alternative hypothesis might say that the frequencies are unequal—meaning either that experimental has higher frequency than control or vice versa. This is known as a *two-sided* alternative. There are two possible *one-sided* alternatives. The first would

be that AE frequency in the experimental treatment is greater than that for the control treatment. The second one-sided alternative would be the opposite of this. In safety monitoring of clinical trials DMCs would normally be interested in the former—experimental greater than control. However, under some circumstances a two-sided alternative might be preferred. The p-value is the probability that we reject the null hypotheses (i.e., declare a treatment difference) when the null hypothesis is, in fact, true (i.e., in truth there is not a treatment difference). In designing clinical trials on the basis of efficacy endpoints we usually set our error rate to be no greater than 0.05. This *Type I error* setting is called the *significance level* of the test and is usually denoted by the Greek letter alpha, α.

For the rofecoxib thrombotic event incidence data in Table 5.2, one method for computing the p-value for the difference in incidence (rofecoxib 3.57%, placebo 2.00%) is the *chi-square test*. The details of computing the chi-square test are contained in most statistical textbooks (see, for example, Rosner, 2006). The p-value computed for the incidence data by chi-square is $p = 0.008$. This means that the probability that this difference could have arisen by chance, if, in fact, incidence between rofecoxib and placebo were equal, is 0.008. Another method of calculating the p-value is by Fisher's Exact Test (Rosner, 2006). This test is implemented in many statistical software packages (see, for example, SAS Institute, 2008; Cytel, 2007). For the data at hand the one-sided p-value calculated by this method is 0.01 while the two-sided is 0.015. A standard criterion for defining statistical significance is p-value less than 0.05 (known as the *significance level*). Both of our methods yield p-values of less than 0.05. Hence, we conclude that the difference in thrombotic AE incidence is statistically significant. This does not mean that the difference is clinically significant. It merely means that the difference that we observed was not likely to have occurred due to chance. Some clinicians, acknowledging that the difference was unlikely to have occurred by chance, might feel that an incidence difference of 1.57 is not clinically meaningful. Indeed, clinicians might see a difference in incidence that they deem clinically meaningful even though it is not statistically significant. We will deal with that situation later in this chapter. For now it is sufficient to remember that if the p-value were calculated to be greater than 0.05, we could not conclude that thrombotic event incidence is equal between treatments. We can only say that we have seen no evidence of a statistically significant difference. A larger sample size might have found the observed difference statistically significant. The results of these statistical hypothesis tests are found in Table 5.3.

It is possible to test the null hypothesis of no difference between Kaplan–Meier time-to-event curves against the alternative of a difference between curves. The p-values for this hypothesis test are generated using the *log rank test* (Cox, 1972; SAS Institute, 2008; Cleves, Gould, Guttierrez et al., 2008). Table 5.3 indicates that the log rank test for our example has yielded a p-value of 0.008 for the treatment difference in time to first occurrence of thrombotic AE. The log rank test requires the proportional hazards assumption. Although this test might often be associated with a primary efficacy endpoint such as time to death, time to disease progression, and so on, this assumption might not be met in safety data. The log rank test should not be the primary hypothesis test for safety data. It should be considered supportive and generated only for those AE types being followed as possible concerns.

TABLE 5.3: Statistical Inference of Adjudicated Thrombotic Adverse Event Data from the APPROVe Trial

	Rofecoxib	Placebo
Incidence (%)	3.57 (46/1287)	2.00 (26/1299)
p-value: chi-square	Chi sq = 5.907, 1 df, $p = 0.008$	
Fisher's Exact Test (one-sided)	$p = 0.010$	
Fisher's Exact Test (two-sided)	$p = 0.015$	
Log rank test	$p = 0.008$	
95% Confidence Interval for incidence (normal approximation)	(2.56, 4.58)	(1.24, 2.76)
95% Confidence Interval for incidence (exact binomial, Clopper–Pearson)	(2.63, 4.74)	(1.31, 2.92)
Incidence rate/100 patient years	1.50	0.78
95% Confidence Interval for rate/100 patient years (normal approximation)	(1.07, 1.93)	(0.48, 1.08)
95% Confidence Interval for rate/100 patient years (binomial approximation)	(1.11, 1.99)	(0.51, 1.14)
Odds Ratio and 95% Confidence Interval	1.82 (1.12, 2.95)	
Poisson Rate Ratio and 95% Confidence Interval (normal approximation)	1.92 (1.19, 3.11)	
Poisson Rate Ratio and 95% Confidence Interval (binomial approximation)	1.92 (1.16, 3.25)	

5.3.3 Confidence Intervals

Confidence intervals (CI) enable us to estimate a plausible range for an unknown parameter. There is a need to compute confidence intervals for our parameter estimates because we must account for the variability of estimates taken from limited samples. The degree of uncertainty decreases with increasing sample size. We now look at various methods of estimating confidence intervals for different parameters encountered in safety analysis. We will introduce some new analytic methods in the process.

5.3.3.1 Incidence

We first look first at confidence intervals for the incidence estimates. There are two methods of estimating confidence intervals that we can use—the normal approximation (Rosner, 2006) and exact binomial method also known as Clopper–Pearson (Clopper and Pearson, 1934; Hollander and Wolfe, 1999). The normal approximation works well when the number of events is greater than 15. For this reason many DMCs prefer the exact binomial method. Table 5.3 shows that rofecoxib thrombotic AE incidence was 3.57%. The 95% confidence interval using the normal approximation, written as (lower limit, upper limit), is (2.56%, 4.58%). The interpretation of this interval is that if we were to perform this clinical trial an infinite number of times and estimated incidence and the confidence interval at the end of each trial as we have, 95% of intervals so calculated would contain the true incidence. Thus we have "95% confidence" in this interval. It is important to note that the frequentist confidence interval definition does not imply that the probability is 95% and that the true value of the parameter lies between 2.56% and 4.58%. In fact, there is nothing magical about these two numbers. They are just estimates that came from one realization of a clinical trial that will be performed only once but, for probability theory purposes, we assume could be repeated infinitely many times. DMCs concerned about the rate of thrombotic AEs would see that a rate as high as 4.58% is within the plausible range implied by the data.

The details of computing the exact binomial confidence limit are not presented here but are presented in the references cited above. The calculations can be made with readily available software (SAS Institute, 2008; Cytel, 2007). The exact binomial 95% confidence interval for rofecoxib is (2.63%, 4.74%) which is pretty close to the normal approximation because the number of events was sufficiently large. Still the binomial interval is asymmetric about the point estimate of incidence while the normal approximation interval is necessarily symmetric. Some statisticians like the elegance of the asymmetric interval. For placebo the corresponding 95% confidence intervals are (1.24%, 2.76%) and (1.31%, 2.92%) respectively. Thus for placebo we have a high level of confidence that the thrombotic AE rate is less than 3.00%. This is further evidence of the separation of distributions of rofecoxib and placebo. It is consistent with the statistically significant difference we found earlier.

5.3.3.2 Rate per 100 Patient Years

If incidence is constant over time, the rate per 100 patient years follows the Poisson distribution (Rosner, 2006). There are two methods for computing approximate 95% confidence intervals for the rate per 100 patient years—normal and binomial. For

the normal distribution method the standard error of the rate per 100 patient years is calculated

$$S = (\sqrt{X}/T) \times 100 \tag{5.1}$$

where X is no. of events and T is patient years. If R equals rate per 100 patient yrs, then the 95% CI is

$$(R - 1.96 * S, R + 1.96 * S) \tag{5.2}$$

In our example for rofecoxib, $R = 1.50$, $X = 46$, and $T = 3059$. Hence $S = 0.22$ and the 95% CI for rate per 100 patient years is (1.07, 1.93).

The binomial method consists of taking the endpoints of the exact binomial confidence limits for incidence and multiplying them by N/T where N is the sample size. For rofecoxib $N/T = 1287/3059 = 0.42$. Hence the binomial approximation for the confidence interval for rate per 100 patient years for rofecoxib would be 0.42×2.63 and 0.42×4.74 or (1.11, 1.99). For placebo the multiplier is 0.39 (1299/3327) and the resulting 95% confidence interval is (0.51, 1.14).

Similarly for placebo the 95% confidence interval would be (0.48, 1.08). These confidence intervals show separation of rates between treatment groups. Our study of confidence intervals leads us to methods of estimating relative risk of adverse events—odds ratio and the Poisson rate ratio.

5.3.3.3 Odds Ratio

The odds ratio is a useful measure of association. In our case it would measure the degree of association of thrombotic AEs with a treatment group. The odds ratio is calculated as follows:

$$C = \frac{P1(1 - P2)}{P2(1 - P1)} \tag{5.3}$$

where $P1$ is proportional incidence in experimental group and $P2$ is proportional incidence in control group.

An odds ratio of C equal to 1 would indicate no association (i.e., null hypothesis of no difference in AE incidence between treatments). The more the odds ratio is greater than unity, the more AE is associated with rofecoxib. An odds ratio less than 1 would indicate that the AE occurs more frequently with placebo. We see in Table 5.3 that the odds ratio in our example is calculated as 1.82 and the 95% confidence interval is (1.12, 2.95). This interval, which does not include 1, is another way of indicating that the null hypothesis is rejected in favor of greater AE incidence in rofecoxib. There are several methods for computing confidence intervals for the odds ratio, and validated software is available (Gart, 1970, 1971; Agresti, 1992; SAS Institute, 2008; Cytel, 2007).

When AE tables are produced for DMC review, it is common for members to glance down the AE list looking for large odds ratios, say greater than 5. Different DMC members will have different cutoff numbers, and the cutoff will depend on the nature of the disease. The purpose is to find treatment differences of interest for further discussion and follow-up. Large odds ratios that are not statistically significant are still

worth noting because this may be an early signal. Some DMC members may prefer a combination of odds ratio cutoff and significance level—such as $p \leq 0.10$ and odds ratio greater than 3. Many AEs associated with large odds ratios will not be of clinical concern. The confidence intervals will shrink with increasing sample size as the trial progresses. The odds ratio may shift downward in time, but trends are worth watching if the AEs are clinically important. When trends are spotted in this manner, the closed meeting minutes should note them so that members will remember to follow up at future meetings. When the DMC is masked to treatment and receives tables classified only by coded treatment groups, the control group will be unknown. In these cases a convention can be established such that the odds ratio is always computed with treatment A as the denominator treatment. In these cases DMC members inspecting tables might look for either odds ratios greater than 5 or less than 0.20 in an initial screening.

Appendix Table A.1 presents adverse events observed for selected marketed drugs in placebo-controlled trials. Exact binomial 95% confidence intervals are calculated for drug and placebo incidence. The odds ratio is presented along with its 95% confidence interval. In going over the odds ratios on this list we can get an idea of how some DMCs might have reacted to safety listings at data review meetings. For seasonal allergic rhinitis drug fexofenadine, a DMC might be concerned with the odds ratio of 5 for both dysmenorrhea and drowsiness would also have to consider that these AEs occur at only 1.5% incidence on fexofenadine and the upper limit of the 95% confidence interval is only 2.7%. Conversely, the odds ratio of 11.3 for the adverse event of flushing for erectile dysfunction drug sildenafil might be of greater concern because sildenafil incidence is 10% reaching a 95% upper limit of 12.3%. The various adverse events listed for fibromyalgia drug pregabalin would, presumably, have generated much discussion for a DMC. Odds ratios for the selected AEs range from 4.9 to 13.6. The odds ratio of 8.3 for dizziness was associated with a 45.0% incidence in the pregabalin group with an upper limit of the 95% confidence interval of 49.1%.

5.3.3.4 Poisson Rate Ratio

The Poisson rate ratio is the ratio of events per 100 patient years in the two groups. In the thrombotic AE example in Table 5.2 the rate ratio equals 1.92 (1.50/0.78). This means that the risk per unit time is 1.92 greater for a rofecoxib patient than for a placebo patient. There are two methods for computing 95% confidence intervals for the rate ratio—normal approximation and binomial approximation.

Methods for the normal approximation were derived by Ng and Tang (2005). We will be using the method they refer to as W3:

$$\begin{aligned} R &= \frac{X_1/t_1}{X_o/t_o} \\ Q &= \ln(R) \\ SE(Q) &= \sqrt{(1/X_o) + (1/X_1)} \end{aligned} \tag{5.4}$$

where R equals Poisson rate ratio, ln equals natural log function, Q equals natural log of R, SE(Q) equals standard error of Q, X_1 equals number of events in experimental group or 0.5 if zero events, X_o equals number of events in control group or 0.5 if zero

events, t_1 equals patient years in experimental group, and t_o equals patient years in control group.

Now compute

$$\begin{aligned} U1 &= Q - 1.96 x SE(Q) \\ U2 &= Q + 1.96 x SE(Q)\,. \\ \text{The 95\% CI for } R &= (\exp(U1), \exp(U2)) \end{aligned} \tag{5.5}$$

where exp is the exponential function. The resulting 95% confidence interval for the rate ratio of 1.92 is (1.19, 3.11).

The binomial approximation to the CI for the Poisson rate ratio conditions on the total number of events in each group assumes that we fixed this number in advance. It creates the proportion

$$P = \frac{X_1}{(X_1 + X_0)}$$

where X_1 equals number of events in the rofecoxib group and X_o equals number of events in control group. Hence, in our case, P represents the proportion of thrombotic AEs that occurred in the rofecoxib group, which is equal to 0.639 (46/72). This proportion has an exact binomial confidence interval of ($P1$, $P2$) or (0.517, 0.749). The 95% CI for the Poisson rate ratio estimated by this method is equal to ($F1$, $F2$) where

$$\begin{aligned} F1 &= \frac{P1(d)}{1 - P1} \\ F2 &= \frac{P2(d)}{1 - P2} \end{aligned}$$

where $d = t_o/t_1$ with t_o and t_1 defined as above. In our example the 95% confidence interval is (1.16, 3.25). The binomial method would be preferred when the number of events is less than 50.

Both methods of computing 95% confidence intervals for the thrombotic event rate ratio yield intervals that exclude 1, leading to the conclusion of a statistically significant difference in rate ratio between treatment groups at significance level 0.05.

In practice DMC members should glance at the listings of these AE rate ratios in a manner similar to that described for odds ratios above. The difference between these rate ratios and the odds ratios is that the rate ratios take exposure into account. This may be important when the exposure differs between the two treatment groups.

Closely related to the Poisson rate ratio is the *hazard ratio* associated with the Kaplan–Meier time-to-event graphs (Cox, 1972; SAS Institute, 2008; Cleves, Gould, Guttierrez et al., 2008). This ratio indicates the relative risk per unit time of a patient in one group having an event compared with another group. As for hypothesis testing, its use requires assumptions that might not be met in practice. One of these assumptions, proportional hazards, was not met for the adjudicated thrombotic AE data in APPROVe in our example (Bresalier, Sandler, Quan et al., 2005) and, hence, the hazard ratio was not calculated. The Poisson rate ratio and its confidence interval are generally sufficient as a time-corrected relative risk measure for routine safety analysis.

5.3.3.5 Inference with Kaplan–Meier Landmark Estimates of Incidence

Confidence intervals can be placed on the Kaplan–Meier landmark incidence using standard errors given in the Kaplan–Meier life tables and referencing the normal distribution. Similarly hypotheses can be tested using normal distribution theory. Common methods are presented in Kaplan and Meier (1958), Rosner (2006), SAS Institute (2008), and Cleves, Gould, Guttierrez et al. (2008). Klein, Logan, Harhoff et al. (2007) present and compare several new methods based on fewer assumptions than the traditional methods.

5.3.4 Data Analysis without Statistics

For certain DMC deliberations use of statistical methods might be considered overkill or even misleading. One example arises often in multinational trials where the DMC is interested in whether AEs are being reported more frequently by investigators in one center, country, or region than in others. Disparities may be due to underperformance or cultural/medical practice differences. Reading tables organized geographically would usually be sufficient for this purpose. The DAC could compute rate ratios or odds ratios comparing regions upon request from the DMC, but because the requested analysis arises only because a large AE incidence has been observed in one region and DMC members can find another region with a low incidence, the *p*-value or confidence interval computed would be suspect because the values compared were selected on the basis of their large difference. In any case, regardless of statistical significance, the DMC will probably want to investigate the origins and either conclude that the differences are cultural or suggest ways of bringing all investigators to a common method of medical practice (use of diagnostics, anti-infective drugs, sterilization, etc.). Inspection of tables of use of concomitant medications and laboratory procedures for these different regions would help in this investigation and could be accomplished without statistical analysis.

5.4 Power

It is important for DMC members to understand the concept of *power*. This concept is most often used in computing the sample size for the trial on the basis of presumed data on the primary efficacy endpoint. However, it is important for those reviewing safety data over time to take this concept into account. Briefly stated, power is the frequentist probability of rejecting the null hypothesis of no treatment group difference when we really want to reject this hypothesis. If a placebo group has a cardiotoxicity rate of 5% and the true corresponding rate on the experimental treatment group is 10% or more, this difference would be considered clinically significant and it would be important to detect this difference in a trial (i.e., reject the null hypothesis, compute a *p*-value of less than 0.05). We would like the probability or the power of detecting this magnitude difference to be high, around 0.80 at least. Because of patient-to-patient variability the smaller the sample size (number of patients in each group) the less likely we are

to detect this magnitude of difference even if it really exists. At the end of the trial there may well be enough patients to have adequate power for this hypothesis test but DMCs are looking at safety data throughout the trial. Biostatistical members of DMCs will want to ensure that the DMC does not conclude at a regular interim data review meeting that they can dismiss cardiotoxicity as an issue for this trial because the incidences of cardiotoxicity reported were placebo 3%, experimental 8% and the p-value was 0.23. This failure to find statistical significance in the difference could be due merely to small sample size. This is why it is correct to say only that no evidence of a statistically significant difference has been found, leaving open the possibility of a different conclusion with a larger sample size.

We return to the thrombotic adverse event data in Table 5.2. Let us assume that the observed placebo rate was the expected true rate of thrombotic events for this population. With a sample size of approximately 1290 patients in both the rofecoxib group and the placebo group, the trial would have the power to detect a rate of 3.6% in the rofecoxib group (this was the observed rate) of 0.80. However, at a data review meeting with only 600 patients in each group, the DMC would have a power of 0.80 to detect a 4.6% rate on rofecoxib and with 300 patients in each group, a power of 0.80 to detect a 6.00% rate. These rates are considerably larger than the clinically significant rate of 3.6%. The sample size of 600 patients would have only a power of 0.51 to detect the clinically significant rate of 3.6% in rofecoxib and the 300 patient sample size would yield a power of only 0.32. A summary of the power analysis for rofecoxib versus placebo will be found in Table 5.4.

This analysis of precision of sample size at data review meetings is sometimes called *assay sensitivity* (D'Agostino, Massaro, and Sullivan, 2003). It is important that at the beginning of each closed session, the DAC biostatistician indicate the assay sensitivity, that is, the magnitude of differences from control group incidence, which the current sample size delivers in terms of power, for different levels of control group event rates.

If the control group were an active control rather than placebo, power would have to be calculated on a two-sided basis because we would consider event rates less than control as well as greater than control to be of clinical significance. This can easily be calculated.

A related concept to power is *conditional power*. While power relates to the data at hand conditional power indicates power to detect a difference at a later time point

TABLE 5.4: Power Analysis of Rofecoxib versus Placebo Assuming Thrombotic Event Rate in Placebo Group Is 2.0%

Sample Size in Each Group	Rofecoxib Rate	Power to Detect Rofecoxib Rate in Previous Column
1290	3.6%	0.80
600	4.6%	0.80
	3.6%	0.51
300	6.0%	0.80
	3.6%	0.32

under an assumption of *drift* (trajectory of future data). A DMC might ask the DAC biostatistician, "If the p-value is 0.08 today, what is the probability that we can reject the null hypothesis at the end of the trial?" The probability referred to is the conditional power. Presumably for an SAE of concern, if the conditional power were high, say 0.90, the DMC might have reason to terminate the trial. Conditional power is calculated most often for efficacy parameters rather than safety parameters, but nevertheless, it is a useful tool available to DMCs. A popular method of computing conditional power is the B-value method (Lan and Wittes, 1988; Proschan, Lan, and Wittes, 2006).

5.5 Multiplicity

At data review meetings the DMC is presented with long lists of AEs for many MedDRA preferred terms within body systems. As was described above, p-values and odds ratios can be computed for each. Recall that the definition of the p-value is the probability that the observed treatment difference in AEs would have occurred due to chance if there was truly no treatment difference. In a list of 100 AEs we would, on the average, expect 5 to be statistically significant due to chance. If chance is solely responsible for the difference, we call this result a *false positive.* DMC members who make lists of AEs with small p-values (say less than 0.10) prior to the meeting will have some false positives on this list. However, further consideration of the severity of the AE, likely relationship to the disease, and so on will often eliminate many of these AE types from concern. Indeed, there will be AE types that are serious and unexpected that have p-values greater than 0.10 that may become the focus of discussion at the DMC closed session. The multiplicity issue does not appear to be a major problem for DMC review of AE lists because statistical significance is not a major factor in selecting AEs for concern.

In some cases there may be interest in reducing the list of AEs worthy of further attention based solely on the statistical significance criterion. We could lower our p-value cutoff to, say, 0.001. Then there will be fewer false positives, but we run the risk of increasing the rate of *false negatives*, that is, those AEs that we think we can ignore because they were not statistically significant by our newly selected stringent requirement (0.001) but are in fact treatment related. Mehrotra and Heyse (2004) have presented an approach to controlling multiplicity in safety analysis using the concept of the *false discovery rate* (FDR). The FDR was first described by Benjamini and Hochberg (1995) as a means of controlling multiplicity. The technique has recently been generalized by Pounds and Cheng (2006). The cells of Table 5.5 represent frequencies of treatment relatedness and declarations of statistical significance that might arise from a DMC safety review table of AEs. Of course we could never construct such a table because we do not know the "truth" about relatedness. In the table we see that the *familywise error rate* (FWER) is the proportion of all hypothesis tests that are true (no treatment effect) but are nevertheless rejected

TABLE 5.5: Definition of the False Discovery Rate (FDR)

Truth of Hypothesis	Declared Not Statistically Significant	Declared Statistically Significant	Total
AE not treatment related	A	B	n
AE is treatment related	Y	Z	N−n
Total	N−R	R	N

Note: FWER = familywise error rate = expected value of B/N (Type I error), FDR = false discovery rate = expected value of B/R.

(declared statistically significant). This is the familiar Type I error. The philosophy of the FDR is that restricting FWER is too conservative. Instead we should look at the FDR or the proportion of all hypotheses declared statistically significant where in fact no treatment effect existed.

Mehrotra and Heyes (2004) use a two-step FDR procedure to flag adverse events as statistically significant. They call this procedure DFDR for "Double FDR," a term coined in their personal communication with the late Professor John Tukey. The example given is that of 40 adverse events defined by preferred term within body system in a vaccine clinical trial involving 296 children within the ages of 12 to 18 months. The naive analysis shows four adverse event types to have p-values less than 0.05. The authors describe the details of computation of the FDR-adjusted p-values. The DFDR procedure consists of setting two error rates, α_1 and α_2. First the authors select the minimum p-value within each body system (group leaders). FDR adjustments are made to both these minima and the individual p-values within each body system. AEs are selected for further investigation if both the body system minimum was less than α_1 and the individual AE within body systems passing the first test is less than α_2. Simulation procedures show that setting $\alpha_1 = 0.05$ and $\alpha_2 = 0.10$ yields good results (i.e., FDR of less than 0.10).

FDR adjustment of p-values within a group proceeds as follows:

Suppose there are k AEs in a group, the p-values are arranged in ascending order, and the highest p-value is unadjusted. For all other p-values the jth in ascending order is replaced by the minimum of $(j + 1)$st or (k/j) $\times$ jth (or itself).

In the example cited by Mehrotra and Heyes (2004), in the first level FDR adjustment the 40 AE types are reduced to 8 candidates and second level adjustment reduces the field to 3 finalists.

Classification of preferred terms to body system is somewhat arbitrary and sometimes the same or similar terms appear within several body systems. Figure 5.4 displays another way of stratifying AEs. There are four discrete categories defined by AE seriousness (serious or not serious) and incidence (high or low). The definition of serious is that given by regulatory agencies and described in Chapter 4. Presumably, a high–low determination could be made by the DAC before each data review meeting. Table 5.6 presents data for an example of FDR adjustment using this type of stratification. The table shows the initial p-values and progressive rounds of FDR adjustment using the method described above. We see that 3 of the four original AE types

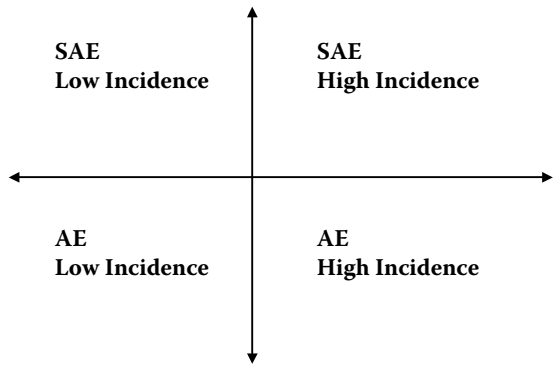


Figure 5.4: Stratification of adverse events by seriousness and incidence.

TABLE 5.6: Example of Use of False Discovery Rate

Group	Adverse Event Type	Initial *p*-value	Group Leader FDR Adjustment ($\alpha_1 = 0.05$)	Within-Group FDR Adjustment ($\alpha_2 = 0.10$)	Meets Both α_1 and α_2 criteria
SAE High Incidence	Anemia	0.02	0.03	0.06	yes
	Renal failure	0.20		0.30	
	Anorexia	0.30		0.30	
SAE Low Incidence	Myocardial infarct	0.03	0.036	0.09	yes
	Dehydration	0.12		0.18	
	Shortness of breath	0.18		0.18	
AE High Incidence	Stomatitis	0.08	0.10	0.10	
	Skin rash	0.10		0.20	
	Fever	0.23		0.28	
	Nose bleed	0.28		0.28	
AE Low Incidence	Diarrhea	0.036	0.048	0.04	yes
	Nausea and vomiting	0.40		0.12	
	Blurred vision	0.26		0.31	
	Bronchitis	0.31		0.47	
	Wheezing	0.40		0.50	
	Headache	0.50		0.50	

(two SAEs and one AE) make the FDR cut. Of course the reduction would likely be more with larger and more realistic tables.

It is important to note that under no circumstances should the FDR adjustment of *p*-values associated with AEs be the sole reason to eliminate any AE type from further discussion. The safety review should be driven by clinical concerns based on the knowledge of the drug and the disease process. The FDR adjustment is a convenient way for the biostatistician to focus discussion in the face of multiplicity.

5.6 Analysis Methods—Likelihood

Likelihood methods of analysis are based on the foundations of statistical inference, but the methods have been developed for applications to biostatistical problems most recently by Royall (1997, 2000) and Blume (2002). Likelihood methodology is a useful tool of inference for DMCs because the methods yield insight on AE incidence, event rates and relative risk conditional on the data we have observed up to the point of data review. Likelihood methods are not based on the repeated sampling foundation of frequentist methods. DMC use of likelihood methods will not get involved with the conundrums of multiplicity. The methods will yield a plausible range of AE incidence or event rates conditional on the data we observed, not, as frequentist methods provide, based on the success rate of a formula including the true value if the clinical trial were repeated infinitely. Advocates of likelihood inference criticize frequentist methods that seek to control Type I (significance level) and Type II (related to power) errors because these errors do not measure the *evidence* gained from the clinical trial for different values of AE incidence or event rates. Only through likelihood methods can relative evidence be assessed. It is relative evidence that is central to DMC deliberations on AE incidence and event rates.

Frequentist methods are *deductive*. To take a clinical example a deductive method would ask: Given the disease what symptoms can we expect? The statistical analogue to this is: Given the parameter of incidence what should the data we collect look like in terms of central tendency and variation? Inference in everyday medical practice is *inductive*. Physicians are generally reasoning: Given the symptoms that are observed in the patient, what diagnoses are most likely, moderately likely, and unlikely? Statistically, given the data we have observed, what parameter values (mean AE incidence, rate per 100 patient years) are most likely, moderately likely, and unlikely?

Consider the toss of a coin. We expect the probability of the coin landing heads to be 0.50. This would be the definition of a "fair" coin. Suppose we toss the coin three times and observe three heads. We now might ask conditional on the data we observed (3 heads) what the relative likelihood is that this coin has the property of probability of head = 0.75, 0.50, and so on. This is the kind of inference we want to use for DMC data review.

Refer again to the incidence of confirmed thrombotic events in the APPROVe trial (Table 5.2). For the rofecoxib patients the event rate per 100 patient years was 1.50.

The reference probability distribution for rate per 100 patient years is the Poisson distribution. The likelihood function may be thought of as reflecting our relative

belief in the magnitude of the parameter (mean rate) given the data collected. The log likelihood function is written as

$$\ln(\lambda|X, T) = -\lambda T + X \ln(\lambda) + C \tag{5.6}$$

where the Greek letter λ (lambda) represents the mean rate, X equals total number of events, and T equals total patient years, and C is a constant term.

From Equation (5.6) we can calculate the *maximum likelihood* (i.e., best supported by the data) value of λ, denoted $\hat{\lambda}$.

$$\hat{\lambda} = X/T \tag{5.7}$$

or the total number of events divided by total patient years. We multiply by 100 in order to express in rate per 100 patient years. We see that the way we have been calculating this rate all along is the maximum likelihood value for this rate from the Poisson distribution.

The value of lambda for the rofecoxib arm of the APPROVe trial that has maximum likelihood is 1.50 (Table 5.2).

We now compute the log likelihood ratio for lambda. It is written as

$$A(\lambda|X, T) = \ln L(\lambda|X, T) - \ln L(\hat{\lambda}|X, T)$$

or Equation (5.6) with $\hat{\lambda}$ substituted in the subtraction term. The likelihood ratio for λ is then

$$LR(\lambda|X, T) = \exp(A(\lambda|X, T)) \tag{5.8}$$

where exp denotes the exponential function.

In other words, this describes the ratio of the likelihood of various values of lambda to the value of lambda with maximum likelihood. Figure 5.5 presents a graph of this ratio for various values of lambda. The graph shows the relative likelihood of a range of values of lambda. The graph reaches a maximum value of one at theta equals 1.50,

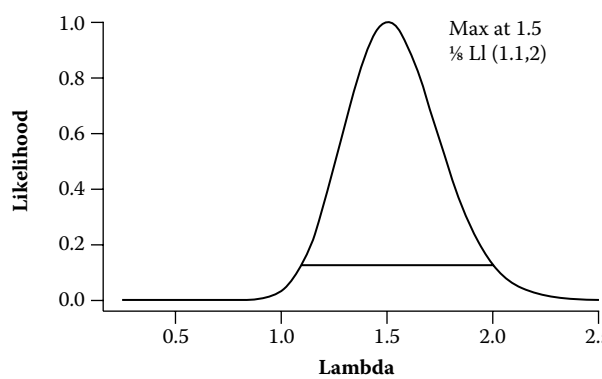


Figure 5.5: Likelihood support graph for rofecoxib group adjudicated thrombotic event rate per 100 patient years (lambda).

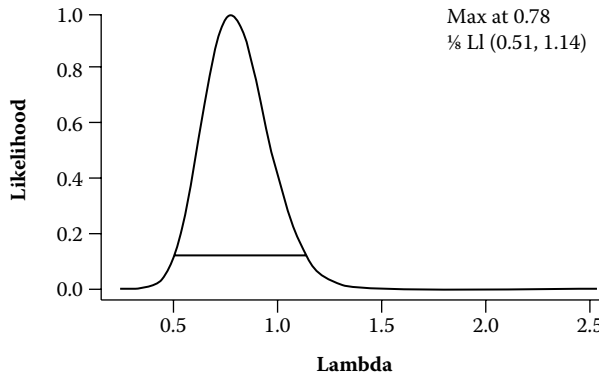


Figure 5.6: Likelihood support graph for placebo group adjudicated thrombotic event rate per 100 patient years (lambda).

the maximum likelihood value. The horizontal line, drawn at 0.125 (1/8), marks the 1/8 *support level*. The line intersects the curve at 1.10 and 2.00. This range, (1.10, 2.00), is called the *support interval* and is somewhat analogous to a 95% confidence interval. It represents a range of values of lambda that are reasonably supported by the data. Royall (1997) and Blume (2002) have indicated that values of theta outside this range are weakly supported by the data collected. If DMC members have a prior concern about thrombotic events on rofecoxib, they might want to review this support interval from meeting to meeting, just to see how the data supports various values of the event rate. For example, from this graph the DMC members would conclude that, although the maximum likelihood value for rofecoxib rate is 1.50 per 100 patient years, the data from APPROVe support values of this rate as high as 2.00, but there is little support for values greater than 2.00. They might also want to review the support curve for the placebo group thrombotic event rate. As Figure 5.6 shows, the 1/8 support interval for placebo rate is (0.51, 1.14).

We now turn to assessing evidence on relative risk or what we have been calling the Poisson rate ratio and denoting it as the Greek letter theta, θ.

$$\theta = \lambda_1/\lambda_o$$

where λ_1 is the event rate in rofecoxib group and λ_o is the event rate in the placebo group.

In order to assess evidence of various values of theta, we must compute the relative likelihood for the ratio of rofecoxib to placebo event rates. To compute this likelihood it is necessary to condition on the total number of events in both groups, that is, regard the total number of events as fixed, as if we specified pretrial that there would be, in this case, a total of 72 thrombotic events at this point in the trial. Under this conditioning the log likelihood of theta reduces to binomial form:

$$\ln L(\theta|X_1, X_o, T_1, T_o, N) = X_1 \ln w + (N - X_1)\ln(1 - w) \qquad (5.9)$$

where X_1 equals total events in rofecoxib group, X_o equals total events in placebo group, T_1 equals total patient years in rofecoxib group, T_o equals total patient years

in placebo group, N equals total number of events $= X_1 + X_o$, and C is a constant.

$$w = \frac{\theta}{\theta + (T_o/T_1)}$$

The maximum likelihood value of theta is

$$\hat{\theta} = \hat{\lambda}_1/\hat{\lambda}_o$$

where $\hat{\lambda}_1$ and $\hat{\lambda}_o$ are defined using Equation (5.7) for each treatment group.

The maximum likelihood value of w is

$$\hat{w} = \frac{\hat{\theta}}{\hat{\theta} + (T_o/T_1)}$$

The log likelihood ratio for theta is

$$B(\theta|X_1, X_o, T_1, T_o, N) = \ln L(\theta|X_1, X_o, T_1, T_o, N) - \ln L(\hat{\theta}|X_1, X_o, T_1, T_o, N)$$

and the likelihood ratio for theta is

$$LR(\theta|X_1, X_o, T_1, T_o, N) = \exp(B(\theta|X_1, X_o, T_1, T_o, N)) \tag{5.10}$$

The likelihood graph for theta is shown in Figure 5.7. Likelihood is maximized at the observed rate ratio of 1.92. The 1/8 support limits are (1.18, 3.22). This interval indicates that conditional on the data collected to date there is weak evidence, or support, for a rate ratio of 1, the value that would represent no difference in risk. This conclusion is the same as that for frequentist methods but the interpretations are different. While the support interval reflects relative evidence conditional on the data observed the frequentist confidence interval reflects confidence in a method of calculation because of its theoretical success rate in repeated sampling. The results of the likelihood analysis of the APPROVe trial event rates are summarized in Table 5.7.

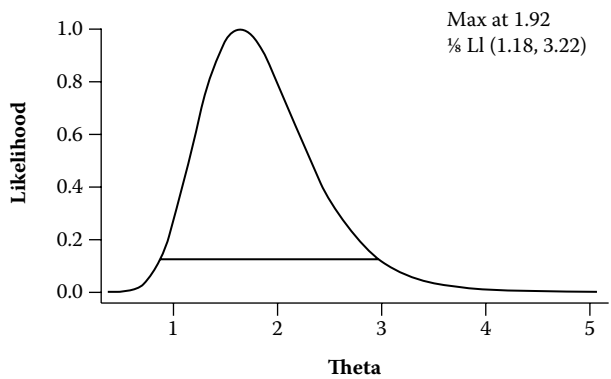


Figure 5.7: Likelihood support graph for the Poisson rate ratio of rofecoxib to placebo thrombotic event rates per 100 patient years (theta).

TABLE 5.7: Likelihood Analysis of Thrombotic Events in the APPROVe Trial

	Rofecoxib	Placebo
Patient years	3059	3327
Number of events	46	26
Event rate/100 pt yrs	1.50	0.78
1/8 support limits	(1.1, 2.0)	(0.51, 1.14)
Poisson rate ratio (theta)	1.92	
1/8 support limits	(1.18, 3.22)	

Table 5.8 compares the confidence intervals and support intervals for the summary statistics of adjudicated thrombotic events from APPROVe. We see that there is considerable agreement in intervals across all methods of interval estimation. Royall (1997) and Blume (2002) would contend that the frequentist confidence intervals have stood the test of time because they approximate the likelihood support intervals which are more consistent with the foundations of statistical inference than frequentist methods.

Frequentist and likelihood methods need not be an either–or. Frequentist methods can be used first. Then, for those AE treatment differences of interest, likelihood methods can be applied while discussions of correction for multiplicity are taking place. If 1/8 support intervals are to be computed for odds ratios or rate ratios for all AE types, then DMC members may wish to concentrate on those AE types whose ratio exceeds some threshold value (say 2) and whose support intervals exclude 1. Just as for frequentist methods, likelihood methods should not be the sole reason to exclude AE types from further investigation.

We have shown likelihood methods of support for the rofecoxib and placebo event rates. We could have also generated these values for the incidence rates for each treatment group and the odds ratio. The likelihood inference for incidence would use the binomial distribution as reference distribution. Royall (1997) presents several methods for computing the support intervals for the odds ratio.

TABLE 5.8: Comparison of Confidence Intervals and Likelihood Support Intervals for Summary Statistics of Adjudicated Adverse Events from the APPROVe Clinical Trial

	Rofecoxib	Placebo
Rate/100 patient years	1.50	0.78
95% CI—normal approximation	(1.07, 1.93)	(0.48, 1.08)
95% –binomial approximation	(1.11, 1.99)	(0.51, 1.14)
1/8 likelihood support limits	(1.10, 2.0)	(0.51, 1.14)
Poisson Rate Ratio (theta)	1.92	
95% CI—normal approximation	(1.19, 3.11)	
95% CI—binomial approximation	(1.16, 3.25)	
1/8 likelihood support limits	(1.18, 3.22)	

5.7 Analysis Methods—Bayesian

A complete description of Bayesian methods as applied to DMC operations in safety review is beyond the scope of this book. The following is a description of the difference between Bayesian methods and frequentist and likelihood methods. Bayesian statistical methods are similar to the likelihood methods described above but extend the inference to incorporate *prior information* on the probability distribution of the parameter of interest such as AE incidence. This prior information might arise from literature review, investigator brochure review, or expert opinion. If no information is available, *noninformative* prior distributions can be specified (Gelman, Carlin, Stern et al., 2004). Bayesian analysis will yield estimates of the probability that the AE incidence is in a range or greater than some threshold value. This range estimate is called the *credible interval* and is analogous to the frequentist confidence interval. However, the confidence interval is not conditional on the data observed or any prior information. It is merely a success rate that the frequentist formula will have in including the parameter of interest in infinite repetitions of the clinical trial. This estimate is based on the *posterior distribution* of the AE inference, which incorporates prior information on incidence and the data collected in the clinical trial. Bayesian methods have been used recently in the design and analysis of clinical trials for efficacy (Berry, 2005). Berry and Berry (2004) have applied Bayesian hierarchical methods to account for multiplicities in safety analysis. This is a Bayesian approach to the FDR methods of Mehrotra and Heyse (2004) described above.

The Bayesian analogue to conditional power is *predictive power* (Proschan, Lan, and Wittes, 2006). The latter is the posterior probability that the null hypothesis of no treatment difference in AE incidence will be rejected at the end of the trial. Its calculation is conditional on the data collected to date and the prior information. Although DMCs might be tempted to use predictive power when a safety concern arises, there are many potential biases inherent in choosing a prior distribution when the treatment difference in incidence at the time of the data review meeting is already known. Lan, Hu, and Proschan (2009) describe the relationship between conditional and predictive power.

5.8 Conclusion

We have described several quantitative methods for analysis of safety data. Not all of these methods will be used during the lifetime of a single DMC. The biostatistician member can work with the other DMC members and the DAC biostatistician to work out what types of data displays and statistical methods are useful for the trial at hand. Indeed the nature of data tables and statistical methods will vary over the lifetime of the trial as safety issues come forward and are put to rest. The statistical methods guide the DMC in their deliberations but do not provide reasons to eliminate safety concerns from further discussion. A summary of statistical methods described in this chapter will be found in Table 5.9. There are many choices of statistical methods

TABLE 5.9: Summary of Statistical Methods for Safety Analysis

Method	DMC Use	Advantages	Disadvantages
Incidence	To review occurrence of AEs	Easy to calculate and understand	Does not take drug exposure (time on treatment) into account
Incidence rate/100 patient years	To review occurrence of AEs per unit time	Takes both occurrence and drug exposure into account	Denominator confusing to some; does not relate to the time on treatment of any particular patient
Kaplan–Meier time to event graph	To graphically review the occurrence of AEs over time since treatment start	Can identify early and late risks; software readily available	Due to limitations of space not practical to be produced for every AE type
Kaplan–Meier estimate of incidence at a time point post-treatment start	To review occurrence of AEs within a time period; one time point on a Kaplan–Meier graph	To identify patient risk within a period of time.	Tendency for some people to create a crude rate by using an inappropriate denominator
p-value	To assess probability that treatment differences in incidence could have occurred by chance if the two treatments in fact had the same incidence rates	All DMC members are familiar with the concept; software readily available.	Often confused with clinical significance
Fisher's exact test, chi-square test	To compute *p*-value for the treatment difference in incidence	All DMC members are familiar with the concept; software readily available	Chi-square test not appropriate for small sample size

(*Continued*)

TABLE 5.9: Summary of Statistical Methods for Safety Analysis (Continued)

Method	DMC Use	Advantages	Disadvantages
Log rank test	To compute *p*-value for the treatment difference in the Kaplan–Meier time to event graph	Allows for statistical assessment of the difference between curves; software readily available.	Assumption of noninformative censoring often not met for safety data; optimal only if proportional hazards assumption is satisfied
Confidence interval	To compute a range of plausible values for a parameter such as AE incidence	Allows DMC members to take sampling variation into account while assessing an AE parameter and to see if a value of interest is within the plausible range	Precise definition often not understood. Many think the interval represents the probability of the parameter, lying in the interval;
Odds ratio	To assess the relative risk of AE incidence between two treatments	Allows for comparison of risks; software readily available	Computation intuitive; many would prefer just to take the ratio of incidences; arbitrary imputation needed when no events occurred in at least one group.
Poisson rate ratio	To assess the relative risk of AE incidence between two treatments taking drug exposure into account	Allows for comparison of risks per unit time	Many different ways to compute the confidence interval for the rate ratio; software not readily available.
Power	To assess assay sensitivity, to see how likely the sample size is to detect differences between treatments	Allows for consideration of why a statistically significant difference between treatments has not been observed	Usually separate software from that used for *p*-value calculation is needed; not practical to compute for all treatment comparisons

Conditional power	To assess the power for detecting a treatment difference at a future point in time	Allows DMC to consider if a treatment difference is likely to be statistically significant at a future data review meeting	Difficult for many to understand; makes many assumptions; software not readily available.
False discovery rate (FDR)	To adjust *p*-values for multiplicity	Can reduce the number of AEs to consider if *p*-value were the chief method of prioritizing AE types for discussion	Not easily understood; may overemphasize role of *p*-values in safety review; stratification methods may be seen as arbitrary; software not readily available
Likelihood or support interval	To compute a range of plausible values of the parameter of interest conditional on the data observed	Avoids the repeated sampling interpretation of the confidence interval	Not easily understood; software not readily available
Bayesian credible interval	To compute the probability of the parameter of interest falling inside a range conditional on the data observed and on prior information	Avoids the repeated sampling interpretation of the confidence interval and incorporates prior information from previous trials or expert opinion	Not easily understood; software not readily available

TABLE 5.10: Statistical Methods to Discuss with the DAC Prior to the Start of the Clinical Trial

1. Patient years of follow-up by treatment group
2. AE Incidence rate, rate per 100 patient years or both
3. Kaplan–Meier time to event graph for selected AE types
4. Kaplan–Meier landmark estimate for all AE types, selected AE types
5. Chi-square test or Fisher's exact test
6. Include the log rank test with Kaplan–Meier graphs
7. Odds ratio, Poisson rate ratio or both
8. Confidence intervals—with point estimates and how to be computed
9. Power/assay sensitivity analysis for selected levels of AE incidence (as in Table 5.4)
10. Whether DAC should be prepared to calculate
    a. Conditional power analysis
    b. False Discovery Rate
    c. Likelihood or support intervals
    d. Bayesian credible intervals

and the DMC should discuss their statistical requirements with the DAC at or shortly after the orientation meeting. Table 5.10 is a suggested list of choices of statistical methods that the DMC should discuss with the DAC prior to the start of the clinical trial. Our next chapter will build on our clinical and statistical knowledge to make DMC members aware of potential biases and inferential pitfalls.

DMCounselor

Q5.1 I am a DMC chair. Our members have been concerned that the sponsor is sending us voluminous tables to review. There are so many pages that present AEs in different ways and with maximum granularity with preferred terms. We are concerned that this information overload might mask true safety concerns.

A Your DMC has come up with an important issue. The sponsor must understand that it is the DMC that decides on the formats of reports. When huge amounts of data are issued to the DMC, it is usually because the sponsor is just dumping the same tables to the DMC that it will later submit for regulatory approval. The sponsor must be prepared to generate separate table formats for the DMC. While we are on this topic, it is useful to note that some regulatory agencies mandate that submissions for licensure break down all efficacy and safety tables by age, gender and race. The DMC should decide at the beginning of the trials with wide eligibility requirements if they find it useful to receive tables stratified in this way.

Q5.2 I am a physician and chair of a DMC. I have inspected the Kaplan–Meier time-to-event curves for a neurological adverse event of concern. The curves cross at the end so I have concluded there is no difference. Our biostatistician member ignores my dismissal of

this as an important event. She keeps talking about early differences. Can you explain what she is talking about?

A The important phrase of your question is "at the end." The curves will often cross at late times because there are very few patients being followed that far and the estimates are less precise than those earlier in time. Your biostatistical member is probably directing the DMC to a landmark analysis (say, 6 months post-treatment start) where she sees some evidence of a difference. This may be an event that occurs early if it is going to occur at all. She is trying to motivate discussion of this. Concentrate on estimates where the data give precise estimates as your biostatistician is correctly saying and don't worry about the tails.

Q5.3 I am the biostatistical member of a DMC working on a trial for an Infant Pharma sponsor. We have been appalled by the extremely poor quality of the data that we are receiving. The sponsor's CRO is performing poorly in monitoring, patient narratives are not arriving in a timely manner, AE forms are incomplete, and we are even concerned that the treatment group assignments may not be correct. DMC morale is low. Can we all just resign from this mess?

A The situation is clearly not acceptable, but you must hesitate to resign. The DMC is responsible for the stewardship of the trial. You are there to protect patient safety and you must consider how this important function will be fulfilled if you were all to exit. The DMC members must bluntly describe expectations for data quality and indicate that your expectations are no lower than those of the regulatory agencies who will ultimately review the data. You may want to provide deadlines for improvement of certain processes, and you may want to go as far as recommending that a new CRO be retained. If there is no improvement, then the DMC should recommend trial termination.

Q5.4 I am the biostatistical member of a DMC. At our organizational meeting I requested that odds ratios and their confidence intervals be computed for all treatment comparisons in safety. We just completed our second data review meeting and the DAC is merely providing p-values for the treatment difference despite my repeated requests. The physician members of the DMC show no interest in this matter. Should I just be satisfied with what I am getting or should I continue to press for what I think would help me and the DMC?

A Once again, the DAC does not specify the data presentation or analysis for a DMC. This is the DMC's responsibility. As the biostatistical member it is your responsibility to see that the analysis is done in the way you think best, and your fellow DMC members should support you on this just as you would support a physician member who has requested a more detailed analysis of QTc interval. You and the chair of your DMC should approach the sponsor and tell them that you are expecting the odds ratios to appear in the tables. The sponsor should investigate what the problem is in meeting your request and get back to you with a plan for resolution. This problem can likely be avoided by an agreement at the beginning of the trial that the DMC must approve all report templates before DAC programming can commence. Also as the biostatistical member you should ask to see draft tables about three weeks before each meeting to make sure that they meet the requirements. There should be no surprises regarding table layout at the time of the meeting.

Q5.5 I am the DAC biostatistician on a cervical cancer clinical trial where overall survival is the principal efficacy endpoint. At the last DMC meeting I attended it was noted that the new institutions just brought into the trial are having a lot of early deaths but they are occurring equally in both treatment groups. The biostatistical member didn't

say anything. He seems to be satisfied on a safety basis because deaths are equal in both groups, but I know that with a lot of early deaths in both groups the sponsor is losing power for the ultimate hypothesis test on overall survival. I mentioned this to the statistical member after the meeting but he just said that the interim analysis of survival is still six months away. He may not understand what I am talking about. I am not sure that I was supposed to comment in the first place but am unsure what to do now.

A I think you were right to talk to the biostatistical member separately and am sorry that your first attempt to bring this to his attention drew that response. If the sponsor is viewing only pooled data, they have no idea that the early deaths are occurring equally in each group, so you must be careful not to unmask the sponsor. Given that this is an oncology trial, they are probably unmasked anyway. Either you or the DMC chair can certainly tell the sponsor that you are concerned about the performance of the new institutions because of the early deaths. These institutions may be admitting patients who are somehow ineligible but will, nevertheless, be included in the intent-to-treat analysis, and the last thing they need to do is lose power. Even worse, the new centers may be giving inadequate care. The biostatistical member should have the first crack at articulating the power problem. I would suggest you call this person and go over the problem, statistician to statistician. I am sure that he will understand and bring up the power issue while the sponsor is evaluating investigator performance at these new institutions.

Chapter 6

Bias and Pitfalls

Bullets to Remember

- Bias may be defined as an observed treatment difference due to effects other than treatments themselves and/or nonobjective actions in operations or evaluations.
- Two potential sources of bias are knowledge of treatment assignment and differential reporting of events between treatment groups.
- It is preferable for the DMC to be unmasked to treatment assignment.
- Reporting bias can occur when there is early termination due to efficacy or with excessive granularity, that is, numerous and narrowly defined MedDRA preferred terms within a system organ class.
- Under a competing risks scenario patients exiting one arm early due to treatment failure can give the appearance of a positive safety profile.

6.1 What Is Bias?

In Chapter 1 we defined bias on an operational level. More scientifically bias is a systematic difference between treatment groups in a clinical trial due to effects other than the treatments themselves (Pocock, 1983). Piantadosi (2005) defines bias as a systematic (nonrandom) error in the estimate of a treatment difference. He goes on to further define bias as a state of mind based on opinion or perception that predisposes actions or evaluations in a nonobjective manner. It is the DMC's task to separate signal from noise. It will be important to consider bias as a carrier of noise. In this chapter we concentrate on sources of bias that can create distorted treatment differences in safety. There are other pitfalls to interpretation of data which do not meet the precise definition of bias but will be covered in this chapter as well, and the term *bias* will be used to apply to all sources of inferential pitfalls.

6.2 Sources of Bias

Sackett (1979) defines several sources of bias that can enter clinical research. In clinical trials we are familiar with sources of bias such as selection bias, publication bias, observer bias, and early termination bias. In what follows we concentrate on those sources of bias that a DMC might come across in safety monitoring. These biases are of two types—knowledge of treatment assignment and underreporting.

6.3 Knowledge of Treatment Assignment

6.3.1 By Sponsor Staff

If sponsor staff are aware of treatment assignment, they might classify SAEs differently or take some actions that will cause bias. Under most circumstances the sponsor staff working on the trial will be masked to treatment assignment until the end of the trial. The DAC will prepare reports for the DMC using coded treatment groups regardless of the DMC masking policy. At data review meetings sponsor staff will see only pooled treatment group information. Pharmacovigilence staff at the sponsor will normally be unmasked to SAEs, but there will be a "firewall" between them and the staff working on the trial, meaning that the treatment assignments will not be revealed to trial staff. For Infant Pharma CROs will often perform the pharmacovigilence function because all Infant Pharma staff may be involved with the trial. Back in the 1980s sponsors were asking FDA permission to take "administrative looks" at the accumulating data. This would entail an interim look at efficacy and safety data by treatment group. Sponsors claimed that these looks were to plan future trials, plan for building manufacturing facilities, plan training of sales staff, etc. FDA now frowns on this activity especially when performed by Infant Pharma sponsors (O'Neill, 2007).

Even with masking there are some analyses of extremes that can be performed by sponsors to weigh the potential of a statistically significant treatment difference in SAE incidence. For example, in the trial of lapatinib plus capecitabine (L + C) versus capecitabine alone (C) for patients with HER2-positive advanced breast cancer (Geyer, Forster, Lindquist et al., 2006), there were 164 patients randomized to L + C versus 152 randomized to C. Using Fisher's Exact test sponsor staff could determine that SAEs that occur with a frequency of 5/164 on L + C and 0/152 on C would yield a one-sided p-value of 0.037. However, a combination of 4/164 versus 0/152 would yield a one-sided p-value of 0.07. The latter would be consistent with a pooled treatment incidence of 4/316 or 1.3%. Thus if sponsors see SAEs with pooled treatment incidence of 1.3% or greater and are confident that these SAEs are rarely or never associated with treatment by capecitabine alone,

they would know that there is a chance of a statistically significant higher incidence in the combination treatment. For pooled incidence of 1.3% or less a statistically significant difference between treatment groups would be impossible. For this trial the authors report a pooled incidence of 5/316 or 1.6% for both asymptomatic cardiac events and fatal AEs. After doing the calculation described above the sponsor might ask the DMC at a data review meeting if there were a concern about SAEs of these types. It is not clear how much bias is introduced by this process.

If sponsor staff is unmasked, their subconscious mind can cause them to be more aggressive in contacting investigators for "clarification" when the terms *related to study drug* or *unexpected* are used for the experimental drug rather than they would be if these terms were reported for a control subject. Pharmacovigilence staff members are usually unmasked, and they must be careful to avoid this practice. Sponsor staff closer to the trial are usually unmasked for oncology trials and may become unmasked as they observe telltale adverse event patterns. DMC members should feel free to inquire about the possibility of these practices.

Sponsor MedDRA coding practices can also introduce bias. One source is granularity in classification, which is discussed below under bias potential at the analysis level.

6.3.2 By the DMC

Some DMCs will choose to be unmasked to treatment from the start of their trial in case they are called upon to make decisions of risk versus benefit. There is also the feeling that DMCs should always be unmasked because this is necessary for their stewardship and is consistent with the trust put in them by the sponsor. In oncology and epilepsy the *add-on* design is common. In this design Treatment X, an approved drug for the indication, is compared with Treatments X + Y, where Y is an experimental treatment. When this design is used, knowledge of treatment assignment might help separate safety issues between X and Y.

When DMCs are masked to treatment, the treatments will be presented as merely A or B. There are two types of DMC unmasking that can take place during a trial—*de facto* and *deliberate*. By *de facto* we mean treatments easily recognizable by the adverse event patterns that emerge as the trial progresses. Sometimes treatment arms will involve different schedules or types of radiotherapy or surgery thus giving away treatment assignment. Also, if patients randomized to experimental treatment will receive infusions but those randomized to placebo will not undergo infusion or if placebo infusion will consist only of saline then the AE "infusion reactions" can unmask the DMC to treatment group. In this case infusion reactions can be presented only for all treatment groups combined. When 2:1 randomization is used or in oncology trials, which are most often open label, treatment group identity will be obvious. De facto unmasking also affects investigators and patients. Both receive lists in their investigator brochures and informed consent documents of adverse events to expect. If the control group is placebo, de facto unmasking is more likely to take place because the adverse event warnings would likely all apply to the experimental

group. If the control group is active control, the lists of expected AEs would be written to include the union for the two treatments without designating which applies to which treatment. However, investigators would certainly know the differentiation and it would not take long for patients to gain the knowledge. DMC members might want to make sure that the routine AE tables include treatment comparisons for those adverse event types that would have an unmasking potential. For a trial where an active vaccine (injected) is compared with a placebo injection, DMC members might want to see tabulations of redness, induration, and pain at injection site. If there is a big differential in incidence between experimental and placebo groups, DMC might be on guard for unmasking.

Deliberate unmasking refers to a decision by DMC members reviewing adverse events with coded treatments to ask the DAC statistician to unmask them to treatment assignment of the individual patients who experienced those adverse events. More will be said about this procedure in Chapter 7. Although a trial may begin with a DMC masked to treatment identity, the DMC may change their policy to be unmasked at any time during the trial.

In pharmaceutical industry trials there is no evidence that DMC unmasking either de facto or deliberate has contributed any measurable amount of bias to their operations.

6.4 Reporting Bias

Reporting bias refers to the conscious or unconscious over- or underreporting of adverse events. We will see various ways in which this may occur. However, it is not the DMC's job to ensure precise estimates of AE incidence. The DMC is looking for signal among noise. For that goal it is important for DMC members to be aware of reporting bias.

6.4.1 Investigator Level

6.4.1.1 Knowledge of Treatment Assignment

Clearly knowledge of treatment assignment can cause investigators to report AEs with either greater or less frequency or to assign relatedness or severity classifications to AEs differently. They might be very wary of the experimental treatment and over report. When the control group is an active control, the investigators might feel that the safety profile for this drug is well understood and fail to report AEs. For the latter DMC members should consult the package insert for the active control and compare the incidence for the particular AE with that observed in the trial. The biostatistician member might want to use the higher incidence to be conservative. Fay, Huang, and Twum-Danso (2007) have recently reported on a method to test the statistical significance of SAE incidence for a

clinical trial with that of a historical control which might be appropriate for this application. Clearly one can posit alternative scenarios for AE reporting level based on investigator knowledge of treatment assignment. This is why many sponsors recommend masking.

6.4.1.2 Incomplete Follow-Up

When a trial ends for early termination due to efficacy, there may be incomplete follow-up on adverse events because investigators feel that the trial is over, are faced with considerable busy work to close the trial, and must now devote more time to other responsibilities. Similarly, in oncology trials when progression-free survival (PFS) is the primary efficacy endpoint and the trial progresses to the end with a statistically significant treatment effect in PFS, some have observed that further follow-up for overall survival, a secondary endpoint, is sloppy (Temple, 2003). This might result in under reporting of AEs. DMC members should insist on complete follow-up and ask the sponsor to provide evidence that AE reporting and follow-up were complete.

6.4.1.3 Spontaneous versus Solicited Adverse Event Collection

Spontaneous AE collection is performed by the physician or nurse. The patient is asked to recall what events the patient experienced since his or her last visit. When AE data are collected by the solicited manner, the data collector goes down a checklist of potential AEs asking the patient to respond to those experienced. Some investigators may probe patients for adverse events that are expected. This is particularly true for approved active control drugs. Sponsors feel that solicitation is needed when there are long periods between clinic visits and patients may forget AEs that were transient and of short duration. Solicited methods are also advocated by sponsors when potentially embarrassing AEs, such as sexual dysfunction, are expected or when the clinical trial involves a treatment for a psychiatric indication such as dementia where memory is expected to be a problem. Wernicke, Faries, Milton et al. (2005) studied three randomized placebo-controlled trials which employed both methods and compared results. Not surprisingly, reporting rates for AEs were higher in solicited collection than in spontaneous, but spontaneous methods were more effective in distinguishing experimental–placebo differences. DMC members should be aware of bias according to type of data collection and should be aware of which method is being used in the trial for which they have stewardship. This type of bias would not apply to SAEs, which are expected to be always reported due to the interventions that are necessary.

The adverse events reported for osteoporosis drugs risedronate and teriparatide in Appendix Table A.1 are of interest. The placebo incidence of arthralgia for risedronate was 23.7%. For teriparatide the placebo arthralgia rate was 8.4%. Arthralgia might be present as part of the disease process. The difference in placebo incidence between trials might be due to different eligibility requirements between the trials, but it might also be because the risedronate trial used solicited adverse events and the teriparatide used spontaneous reporting. Which method was used in each trial is unknown.

6.4.2 Analysis Level

6.4.2.1 Early Termination Due to Efficacy

Early termination of a trial due to efficacy might also provide bias on the analysis level. If a trial terminates at a planned interim analysis, the DMC can now compare treatment groups for safety and find no important treatment differences. However, the failure to find differences might be because the sample size is not large enough to find a difference. The biostatistician member of the DMC can compute power and assay sensitivity as described in Chapter 5 to provide a quantitative assessment of potential bias. However, the DMC may suggest continuing to enroll patients on the experimental treatment and reaching a reasonable exposure time in order that the drug's safety profile can be best described. The number of patients in the control group will not increase but, depending on the disease and safety issues, may be sufficient for comparison. If desirable, the control group can be expanded by meta-analysis. This method will be described in Chapter 7.

6.4.2.2 Granularity Bias

Granularity refers to the number of MedDRA preferred terms listed in data review tables within a system organ class (SOC). Table 6.1 represents such a table from an infectious disease trial. Here there are 9 preferred terms within the SOC "eye disorders." For the SOC "cardiovascular disorders" there may be 50 preferred terms listed. This granularity is created when the DAC generates tables listing all preferred terms within an SOC that had at least one occurrence in either treatment group. There is concern that excess granularity may hide a signal.

Table 6.1 shows that there is a statistically significant difference in eye disorder incidence overall between treatments A and B by Fisher's Exact Test (one-sided). However, there is not a statistically significant difference for any of the preferred

TABLE 6.1: Incidence of Adverse Events for MedDRA System Order Class Eye Disorders by Treatment Group and Preferred Term

	A ($n = 440$)	**B** ($n = 440$)	***p*-value (one-sided) if $p < 0.05$ by Fisher's Exact Test**
Any event within class	4	17	0.003
By Preferred Term			
1. Chorioretinal disorders	1	0	
2. Conjunctival hemorrhage	1	2	
3. Conjunctival edema	0	2	
4. Conjunctivitis	1	5	
5. Eye inflammation	1	0	
6. Eye edema	0	2	
7. Eye redness	0	2	
8. Keratoconjunctivitis sicca	0	2	
9. Scleral edema	0	2	

TABLE 6.2: Incidence of Adverse Events in Table 6.1 Grouped by Preferred Terms Representing Inflammation Types

Groupings Representing Inflammation Types	A ($n = 440$)	B ($n = 440$)	p-value (one-sided) if $p < 0.05$ by Fisher's Exact Test
1	1	0	
2	1	2	
8	0	2	
3,4,5,6,9	2	11	0.011

TABLE 6.3: Incidence of Adverse Events in Table 6.1 Grouped by Preferred Terms Representing Types of Tissue Involved

Groupings Representing Types of Tissue Involved	A ($n = 440$)	B ($n = 440$)	p-value (one-sided) if $p < 0.05$ by Fisher's Exact Test
1	1	0	
9	0	2	
2,3,4,8	2	11	0.011
5,6,7	1	4	

terms. Suppose this committee had an ophthalmologist member because of expected eye-related AEs. This person might say that the overall treatment difference is difficult to interpret because the preferred terms are heterogeneous. Tables 6.2 and 6.3 show two attempts at grouping the preferred terms to display adverse events by inflammation type and tissue type, respectively. Each of these tables produces one subgrouping with a statistically significant difference.

The problem occurs when DMC members are unmasked either de facto or deliberately. There may be a subconscious attempt to hunt for subgroupings that provide a rationale for the difference when there is general agreement that the overall difference is meaningless because the preferred terms are heterogeneous. The problem is compounded when two specialists disagree on the grouping, and even the same specialist might group preferred terms differently on two different occasions. This scenario can also occur when there is no evidence of a statistically significant difference for the SOC overall, but specialists engage in preferred term clumping to seek statistically significant differences.

Obviously a considerable amount of time can be spent in these discussions, and this activity may cause the DMC to request that representatives of other medical specialties be retained as consultants to give input in combining preferred terms that occur in SOCs within their specialty. Although the latter is rarely done, many DMC members wonder if there is some hidden signal within the noise introduced by granularity in preferred terms.

Goldman (2002) discusses the art and science of MedDRA coding especially pointing out biases that can occur in the process. He points out that the appropriateness of the terms *expected* and *unexpected* can depend on the degree of granularity in

classification. This could result in an SAE being coded as an AE. He worries that the term *congestive heart failure* can be bypassed by merely coding the patient as having *dyspnea*, *orthopea*, and *fatigue*. If an investigator used only two of these three terms, it is more likely that congestive heart failure would not appear as a signal. The degree of sponsor aggressiveness in clarifying this issue with investigators could depend on whether they have knowledge of patient treatment group. Similarly, White (1998) observes that a strategy of intense "splitting" (granularity) can keep SAEs off the label for the product postapproval.

Granularity bias can be minimized if the DMC can agree on preferred term combinations before the start of the trial based on expected AEs, mode of action, and pathways associated with the experimental drug. If there can be agreement as to which AE types would undergo statistical analysis of the sort shown in Tables 6.1 through 6.3 at the start of the trial, the biostatistical member can suggest analyses to control multiplicity problems. As the trial progresses and the safety profile is better understood, less granularity may be needed than was thought necessary at the beginning of the trial. Kubler, Vonk, Belmel et al. (2005) suggest the establishment of minimum incidence required for each level of specificity. For example, for 1% incidence perhaps only the system organ class (SOC) could be presented; for 5% the preferred terms (unique medical concepts) could be displayed. This might ensure that there are enough patients for subclassification.

In the presence of a long list of subclassifications, DMC members might ask if the direction of differences trends toward a certain treatment. In Table 6.1 DMC members might ask if there is a trend in treatment differences in AE incidence in the direction of treatment B. Of 9 subcategories there are 7 in which treatment B has a higher count of eye disorders than treatment A. The biostatistician member of the DMC might want to refer to the binomial distribution with a null hypothesis that the probability of treatment B having a higher count than treatment A for any subcategory is 0.50. Under this null hypothesis, the probability of observing 7 or more differences in the direction of treatment B out of 9 subcategories is 0.09. Thus there is some evidence of a trend, but the difference is not statistically significant at the 0.05 level.

There are no approaches to granularity bias that apply to all contexts but DMC time should be guided by seriousness, severity, relatedness, unexpectedness and should take investigator brochure information on mode of action, pathways, etc., into account. This should not be thought of as strictly a statistical issue.

6.5 Competing Risks

Consider a hypothetical lung cancer clinical trial. The control group is Treatment A, an already approved drug, and the experimental group is Treatment A plus experimental drug B. The primary efficacy endpoint is progression-free survival (PFS) meaning that patients go off study if they die or show evidence of disease progression. Table 6.4 presents data on the frequency of cardiac SAEs from an interim analysis in such a trial. The typical analysis would ignore patient years of follow-up as shown in the first

TABLE 6.4: Cardiac Serious Adverse Events by Treatment Group—Interim Analysis Results from a Hypothetical Clinical Trial for Lung Cancer

	Treatment A $n = 150$ pt yrs = 175	Treatments A + B $n = 150$ pt yrs = 85	Analysis
			$p = 0.018$, 1-sided Fishers Exact Test.
No. Cardiac SAEs	8 (5.33%)	1 (0.67%)	OR[1] = 8.39 95% CI: 1.04–67.97
Cardiac SAEs/ 100 pt yrs	4.57	1.18	RR[2] = 3.89, 95% CI: 0.49, 31.07

[1] OR: Odds Ratio
[2] RR: Relative Risk

line of the table. The incidence of cardiac SAEs for Treatment A was 8/150 (5.33%) and for A + B 1/150 (0.67%). The one-sided Fisher's exact test shows this difference to be statistically significant (p = 0.018). Some DMC members might conclude that this indicates that although Treatment A is associated with cardiac SAEs, Treatment B may have a protective effect on cardiotoxicity.

However, if we look at the patient years of follow-up, we see that the active control has 175 years and the experimental treatment has only 85 years. Further investigation might show that patients on the experimental treatment are exiting the trial sooner due to disease progression or death. Thus the treatment A + B patients are perhaps not exposed to their drug long enough to experience *cardiotoxicity.* This is an example of *competing risks*. Loosely defined, the competing risk phenomenon occurs when a patient can experience several types of event but the occurrence of one (in our case treatment failure) lowers the probability of observing another event (cardiac SAE).

If we include patient years follow-up in our analysis and compute the number of events per 100 patient years, Treatment A has 4.57 versus 1.18 for Treatments A + B. The relative risk (RR) is 3.89 (4.57/1.18). With reference to the Poisson distribution the 95% CI is 0.49, 31.07 which includes 1 and, thus, this difference is not statistically significant. The failure to attain statistical significance is due to the large variance component introduced by the shorter follow-up in the experimental arm. This is an example of why it is important for DMC members to review at every meeting a table of frequency of discontinuation by reason and treatment group and for the DAC to report patient years of follow-up by treatment group. In practice this particular clinical trial might meet the conditions for early termination due to efficacy at the next planned interim analysis, but competing risks should always be addressed as a source of bias in comparing treatment groups for safety.

A more precise safety comparison taking follow-up into account would be a Kaplan–Meier time-to-event analysis. However, neither it nor the Poisson analysis described previously corrects the treatment comparison for competing risks. Gray (1988) presents a method of time-to-event analysis correcting for competing risks. A good overview of competing risk methods is presented by Pintilie (2006).

The analysis performed above was not planned and, as has been emphasized previously, the safety issues that arise are not strictly statistical matters. Despite the competing risks, there might very well be a protective effect of Treatment B on cardiotoxicity and this characteristic would have emerged had patients been followed longer on the experimental arm. Clinical knowledge is very important here. All we can conclude statistically is that there is no evidence of a statistically significant difference. This leaves open other interpretations and conclusions from other trials with longer follow-up.

6.6 Conclusion

We have seen the various sources of bias and how they may be minimized. Other sources of bias may appear during the conduct of a clinical trial and DMC members should be considering this possibility in their deliberations. A summary of types of bias and recommended actions to minimize bias can be found in Table 6.5. In our next chapter we take our knowledge of statistics, clinical issues, and bias to DMC decision making.

DMCounselor

Q6.1 I was asked to serve on a DMC for an Infant Pharma company. The medical director told us at our first meeting that he can be unmasked at any time. The venture capitalists agree with him. Should I serve on this DMC?

A This is not an uncommon occurrence among companies in Infant Pharma. Both the executive staff and the investors are not accustomed to the rigors of good clinical practices but they understand the need to terminate trials early that may never approach statistical significance or where there may be a legal risk due to safety concerns. You ought to explain to them the credibility and statistical problems they will face by being unmasked and refer them to consultants who may help them draw up planned interim analyses. If they insist on being unmasked, then you may have to turn down their offer to serve.

Q6.2 I am a biostatistical member of a DMC for an experimental treatment for acne vulgaris. We are masked to treatment assignment and have no efficacy responsibilities. Treatments are only identified as A, B. Whenever we see an odds ratio for an SAE of 1.5 or greater, one dermatologist member insists on being unmasked regardless of statistical significance. He is not comfortable in assessing risk unless he knows what A, B are. He claims other DMCs he has served on have always unmasked for SAEs.

A It appears that the member is not asking to be unmasked for just the treatments of those patients experiencing the SAEs. He is asking to know what A, B stand for. This member may be overly conservative about SAEs because of the controversy over isotretinoin and depression (Magin and Smith, 2005). It is not clear why this person takes an odds ratio

TABLE 6.5: Summary of Recommended Actions to Reduce Bias

Source	Description	Action
Knowledge of Treatment Assignment		
Sponsor staff	Possible bias in classification of SAEs on experimental treatment	Separate pharmacovigilence unit at sponsor independent of clinical trial staff
DMC members	Bias in interpretation of safety data	Not considered a problem; DMC should be unmasked from start of trial
Investigator	Bias in reporting AEs	Masking preferable, if not compare reported AEs with label for active control; compare treatment groups for those AE types that have an unmasking potential
Incomplete Follow-up		
Early term due to efficacy	Poor safety follow-up	DMC should work with sponsor to ensure complete follow-up
	Lack of power	DMC should calculate assay sensitivity in assessing a treatment difference in safety; perhaps perform meta-analysis
Spontaneous versus solicited adverse event collection	Solicitation of adverse events from checklists can create overestimates	Investigate how solicitation might affect difference between experimental and control groups; should have no effect on SAEs
Granularity	Too many preferred terms under a system organ class makes it difficult to find treatment differences and can hide SAEs.	DMC and sponsor should agree on meaningful preferred terms at start of trial; binomial analysis of trend; minimum incidence requirement for subclassification
Competing risks	Patients exit one treatment early due to treatment failure and thus are not on study long enough to develop AEs	Analyze reasons for discontinuation; report patient years follow-up for each treatment group; Poisson or time-to-event analysis

of 1.5 as a cutoff. In assessing risk when unmasked, DMC members would usually have to have access to efficacy information so that risk versus benefit could be assessed. If one member is uncomfortable being masked, it is probably best for the whole committee to be unmasked taking precautions to avoid bias. Given that all members will now be aware of what treatments A and B stand for, it would also be important to ascertain what process this member will go through in assessing risk in the absence of efficacy data whenever an odds ratio exceeds 1.5 regardless of statistical significance.

Q6.3 I am serving as biostatistical member of a DMC for a pediatric solid tumor clinical trial. Our chair is a pediatrician and a world-renowned expert in pediatric oncology. At a recent meeting she pointed out a treatment difference in renal adverse events and I pointed out that this observed difference could easily be a consequence of competing risks because patients in the arm with a lower level of renal events are also exiting the trial sooner due to treatment failure. Her reply was that we don't see competing risks in children, and she went on to other business. I took the floor again trying to explain that competing risks were not a function of the disease process but a statistical artifact. At this point the other physician members of the panel followed our chair and went on to other business. Should I complain to the sponsor?

A No, don't bring up closed meeting information to the sponsor. This is something that you must work out within the DMC. You could ask the chair to explain why "we don't see competing risks in children." Her reply should indicate to all that she is confusing competing risks with some other phenomenon. In explaining the statistical issue you should mention that you respect her knowledge of pediatric oncology but she must respect your knowledge of statistics. In some diplomatic way you might remind her that statistics is your responsibility. Both clinical and statistical expertise must be represented on this committee. Come prepared with examples of competing risks, perhaps from published clinical trials in pediatric oncology, to illustrate your point. It is not likely that this problem will persist.

Chapter 7

Data Monitoring Committee Decisions

Bullets to Remember

- All DMC decisions are *advisory* to the sponsor.
- The DMC must make safety decisions without knowledge of efficacy.
- Statistical significance is neither a necessary nor sufficient condition for the DMC to take action.
- When an SAE of concern is found, the DMC may unmask, recommend that the sponsor issue a "Dear Investigator" letter, modify informed consent and/or protocol, or terminate the trial.
- Prior to unmask for further consideration of an SAE, the DMC should create a decision matrix showing what action would be taken on the various scenarios of SAE distribution by treatment group.
- In making recommendations short of trial termination, the DMC must be careful not to unmask the sponsor.
- At the final meeting or shortly thereafter, the DMC may review the integrated summary of safety, proposed package insert, and manuscripts prepared for publication of trial results.

7.1 Types of DMC Decisions

We have learned about the DMC's role in a confirmatory trial safety monitoring program, clinical issues, statistical issues, and sources of bias. We now turn to how these issues interact when the DMC has to make decisions. The DMC decision most on the mind of sponsors and DMC members is the recommendation to terminate a trial due to safety or persuasive evidence of efficacy or *futility* (i.e., it is very unlikely that, if continued, this trial would conclude with evidence of efficacy). This is an important decision and we will see what information is available to the DMC to help in this decision. Other decisions involve the need to write a "Dear Investigator" letter informing investigators of potential risk, and recommending changes in the informed consent document. At the end of the trial the DMC might be asked to review the

integrated summary of safety in the regulatory submission, manuscripts presenting results of the trial, and the proposed drug label or package insert.

A quorum is usually sufficient to hold a DMC meeting, but most DMC members agree that before any decision can be considered finalized, all members must be consulted.

It is important to note that all DMC decisions are *advisory*. The sponsor makes all of the final decisions and is cognizant of the risks inherent in not following a DMC recommendation. In practice once a DMC uncovers an issue, they generally negotiate a solution with the sponsor. Rarely, however, would the sponsor refuse to terminate a study after receiving a recommendation from the DMC. An intermediate step might be to suspend enrollment of new patients until more information can be gathered by the sponsor's pharmacovigilence staff and the DMC members.

7.2 Decision-Making Environment

The DMC is making decisions on safety without knowledge of efficacy. Even when DMC members receive some efficacy information during the trial, efficacy is not usually established until the end of the trial. For many non-life-threatening indications DMCs will view SAE evidence differently than they would for a more serious indication.

There are some limitations to considering efficacy data available during the trial. First, there is a lack of power to detect a treatment difference early in the trial. Second, some DMC members are not specialists in the disease but rather in expected adverse events. The AE experts together with the biostatistician may be uncomfortable making efficacy decisions. For some small DMCs there may be only one physician on the panel capable of making the decision. This person or persons will often claim that the efficacy endpoints established for the trial are not the way specialists evaluate patient benefit in practice. Metrics such as the SLEDAI in systemic lupus erythematosus (Bombardier, Gladman, Urowitz et al., 1992), PASI score in psoriasis (Fredriksson and Pettersson, 1978), the Ritchie score in rheumatoid arthritis (Lewis, O'Sullivan, and Rumfield, 1988), and so on were all developed for cohort evaluation but not for evaluating an individual patient, and there are still questions about their appropriateness in clinical trials (Ashcroft, Li Wan Po, Williams et al., 1999). It is true that considerable other patient information may be available but the DMC would have to make a special request to the sponsor for this information and such a request might send a signal to the sponsor that a significant safety issue has arisen.

7.3 Risk versus Benefit Analyses

The sponsor will do some kind of risk versus benefit analysis as part of the regulatory submission. Of course, the objectivity of this analysis can be questioned. The DMC will have limited information to perform such an analysis, but one recent example of

DMC decision making for risk versus benefit is the gradual discovery of cardiovascular risks in the PERT trial. Wittes, Barrett-Connor, Braunwald et al. (2007) present a detailed history of their collaboration and risk versus benefit decision to terminate the trial. In the U.K. the National Institute of Health and Clinical Excellence (NICE) does perform these analyses after drug approval (UK National Health Service, 2008) and the German Institute for Quality and Efficiency in Health Care (IQWiG; 2008) provides similar analyses as part of the German drug approval process. Many factors are involved in such an analysis, and a detailed analysis is beyond the scope of a DMC. Most often risk versus benefit analyses such as those by Siu and Rowinsky (1998) for irinotecan in solid tumors, Ziemssen, Neuhaus, and Hohlfeld (2001) for glatiramir acetate in multiple sclerosis, Mikuls and Moreland (2003) for infliximab in rheumatoid arthritis, and Cranney and Adachi (2005) for raloxifene in postmenopausal osteoporosis are done only after considerable postmarket experience with a drug. No single methodology is employed by these authors or others. Risk versus benefit analysis is an important and evolving field.

7.4 When a Safety Issue Arises

We now come to the point where a DMC has found an SAE that appears to occur more in the experimental group than the control group, and they feel that what they know about efficacy will not overcome this differential. This discussion assumes that the confirmatory trial in question has only one experimental arm. Clearly if there are more than one experimental arm representing different doses, schedules, formulations, and so on, safety actions can be taken on only certain experimental treatment groups, not necessarily on all groups. How does the DMC proceed for a single experimental group? First, statistical significance is neither necessary nor sufficient for a DMC to take action. Statistical significance is a function of power of the statistical tests used, and of course, it is vulnerable to multiplicity and the unplanned use of formal statistical inference for the SAE. The experimental treatment might have only a slightly higher incidence of a life-threatening cardiovascular SAE than the active control, but if the drug will eventually be marketed to hundreds of thousands of patients, it will translate into so many more potential public health problems that it may be unethical to continue to expose the clinical trial subjects to this drug. Of course, the DMC members will be reading case narratives and, perhaps, requesting further information. The discussion around narratives—predisposing conditions, function of age and gender, concomitant medications—will often generate possible alternative explanations for the observed SAEs.

DMCs often face the question of how long they should wait before they inform the sponsor of the safety issue. This depends on the individual circumstances of the SAE, indication and other aspects of the drug's safety profile. A calculation of conditional power as described in Chapter 5 might be useful here. Further discussion of this matter will be found below under "Trial Termination."

TABLE 7.1: Possible Actions to Be Taken by a DMC after Observing an SAE of Concern (Listed in Ascending Order of Trial Impact)

Action	Description	Comments
Unmasking	Identification of treatment groups for entire cohort or just for patients with the SAE	This is only a transaction between DMC and DAC statistician in the closed session. No sponsor involvement.
"Dear Investigator" Letter	DMC recommends that sponsor write a letter to investigators asking them to beware of SAE signs and symptoms and, where appropriate, giving suggestions for prevention of the SAE.	There are certain cultural implications in multinational trials where suggestions for medical practice and/or prevention are given in the letter.
Modification of informed consent	DMC recommends to sponsor that specific wording of the SAE be included in a modified informed consent.	Sponsor can only recommend changes to informed consent to investigators. It is up to the investigators and their IRBs to implement. Reconsenting for all patients is recommended over just reconsenting for new patients.
Protocol modification	DMC recommends that protocol be modified for dose, schedule, eligibility, etc.	DMC must be assertive about the change; sponsor may want to wait until other modifications accumulate and make all changes at once.
Trial termination	After waiting sufficient time for accumulation of further evidence DMC recommends trial termination.	DMC should view this recommendation as the beginning of a dialogue with sponsor. Sponsor may have further information about the SAEs and/or other options.

When a safety issue arises through the DMC's review of safety data, a number of activities can take place ranging from unmasking to trial termination. A summary is presented in ascending order in Table 7.1.

7.4.1 Unmasking

In many cases the DMC will be unmasked either by plan or de facto. In some cases the DMC will be partially unmasked and a decision may be made to identify what the group symbols (A, B or Blue, Green, etc.) represent. Alternatively if there are,

TABLE 7.2: Hypothetical Decision Matrix for Seven Patients Reporting an SAE of Concern

Frequency of SAE by Treatment Group		
A	**B**	**Action**
0	7	Request identity of A, B. If B is experimental, begin discussion of ascending decisions as in Table 7.1. If control is active, discuss if expected. If not, search for more information on active control drug.
1	6	Same as 0–7 above
2	5	Same as 0–7 above
3	4	No action
4	3	No action
5	2	Same as 0–7 above with roles of A, B reversed.
6	1	Same as 5–2 above
7	0	Same as 5–2 above

say, seven patients who have an SAE of some concern and the split is 1 versus 6 or 0 versus 7, the DMC may seek further information including the identification of the groups. This method reveals neither the treatments of the individual patients nor the group identities. Even if the DMC requests treatment assignments of the seven patients they have learned only the assignment of those patients. Many DMCs think all of this is a waste of time and prefer just to be unmasked from the start of the trial. As was said earlier, masking usually begs DMC members to guess treatment assignment, which they usually do correctly, but when wrong it can create havoc with their decisions for the rest of the trial.

If this unmasking scheme is to be pursued, the DMC chair should first lead the committee through a *decision matrix*, which would involve what action the committee would take for each possible outcome of the unmasking. If the committee cannot agree on an action, there may be no reason for unmasking. Table 7.2 indicates possible actions that might be taken in the case of 7 SAEs. Actual actions would depend on the individual trial.

7.4.2 "Dear Investigator" Letter

The DMC may feel that it is sufficient at this point to recommend that the sponsor send a "Dear Investigator" letter to each investigator informing them to watch for signs of the SAE and to inform patients to contact the investigator if they experience symptoms related to this SAE. The letter may also contain suggestions for preventing the SAE if, for example, the SAE was a postsurgical infection. The wording here must be diplomatic because it can be taken by some investigators as insulting. Something like a postsurgical infection in a global trial brings up many differentials in training, practice of medicine and supervision. Although the letter may be aimed at certain countries, it must go to all investigators. A request for a "Dear Investigator" letter

is usually accompanied by a request that the sponsor's clinical research associates enforce the letter through on-site discussion with investigator and staff and follow-up auditing. At future meetings the DMC should inform the DAC that they would like to see a pre/post-letter analysis of change in incidence. If there is not a reduction in incidence as a result of the letter, further action may need to be taken.

The DMC will usually ask the sponsor to draft the "Dear Investigator" letter and submit it for comment by the DMC before sending.

7.4.3 Modification of Informed Consent

The DMC may request a rewording of the informed consent form adding a warning to patients of the SAE of concern. The sponsor may send out a suggested rewording, but each investigator site has the right to word the modification in its own way and negotiate the change with their IRB or ethics committee. When a modification of informed consent occurs, the question of reconsenting patients already in the trial arises. It is not considered good practice to show the modified consent form only to new patients entering the trial. Existing patients may be asked to sign the modified consent form at their next clinic visit. An alternative is to call back all patients, regardless of date of next clinic visit, for their review and signing of the new informed consent document. This should be done only in the most extreme cases where the SAE is life-threatening and the patients are taking oral medication. In protocols where patients are dosed only through injections or infusions at the clinic, calling in patients for review would be unnecessary. Many investigator staffs are communicating with their patients through Web sites or e-mail. Hence it is possible to make patients, or their family members/caregivers, aware of a new SAE prior to their coming into the clinic. It is possible that some patients will refuse to sign the modified informed consent and drop out of the trial. This is their right. That this, or reduced compliance to oral medication, might be a consequence of reconsenting should not be part of the decision of whether or not to reconsent.

7.4.4 Protocol Modification

The DMC may feel that the presence of certain SAEs indicates that a protocol modification is in order—change in dosage/schedule, eligibility, and so on. Sponsors will often reply that, due to the long process of running the revised protocol through all IRBs involved, they prefer to wait until enough changes come about from other sources and make the change all at once. The DMC must be prepared to indicate that the time for change is now if they feel that patients are at risk.

7.4.5 Trial Termination

A recommendation of trial termination usually does not come at the first sign of a disturbing SAE treatment difference. What appears to be a disturbing treatment difference early in the trial could be a random occurrence that will correct itself in time. In time-to-event analyses it could be a manifestation of the "bad news travels first" phenomenon. In this scenario SAEs are reported as soon as they occur but

before reports of follow-up time for the many patients not experiencing these events. Hence, the analysis has an overrepresentation of events compared to exposure. The DMC may request another analysis of an appropriate subset of the routine analysis before the next scheduled meeting. This ad hoc meeting will usually be covered by teleconference but, in some cases, the DMC may feel it necessary to meet face-to-face to go over enhanced and updated narratives and other information apart from statistical reports that the DMC may have requested previously. It is important for the DMC to take turnaround time into account when requesting ad hoc reports from the DAC (see *DMCounselor*, Q7.5 below)

If the DMC feels that trial termination is a viable option, they should take this as the beginning of a dialogue with the sponsor where the DMC presents this as their pending decision. This conversation would usually be with the sponsor's pharmacovigilence unit rather than the protocol team. This will keep the latter properly masked while discussions proceed. As was mentioned above the pharmacovigilence unit might first offer to stop enrollment of new patients on the trial while discussions and analysis continue. The pharmacovigilence staff may have information about the drug or the SAEs that the DMC has not considered. The DMC should not be intimidated by responses from the sponsor but should definitely listen to the points being made.

The sponsor will ask the DMC if they have considered the overall benefit as well as risk. The issues involved in risk versus benefit have already been discussed. The relevance of taking that matter up will depend on the indication and the nature of the SAE. In oncology and congestive heart failure trials, more risk may be taken if there is some evidence of benefit or at least a feeling that there is no reason to doubt the hypothesized benefit would outweigh the harm. In trials of seasonal rhinitis or skin rash there may be less tolerance for SAEs especially with effective drugs already on the market with a more favorable safety profile.

For many indications, deaths and SAEs may be considered by the sponsor to be part of the disease process. If there is an excess of these events in the control group, the sponsor may argue that this is not a safety issue at all but early evidence of efficacy. The sponsor may hesitate to terminate the trial early because of uncertainty if the regulatory agencies would accept the curtailed trial as evidence worthy of drug approval. On ethical grounds the sponsor may indicate that the control group patients are at no greater risk than any person who has the disease under investigation not participating in the trial. While this defense is reasonable, the DMC may want the sponsor to at least set a stopping rule by which they would indicate how much of an imbalance would not be acceptable. This stopping rule would have to be discussed with the regulatory agencies to determine how the trial could be salvaged if it had to be terminated according to this rule. In cases such as these and similar, it should be clinical and ethical issues that determine the course of action not the result of a statistical hypothesis test.

The DMC must understand the financial implications of their decision and expect there to be tension in these deliberations. In certain regions of the world the risk versus benefit decision in the presence of an SAE may be different than would be the case in North America or western Europe, and DMCs should be aware of these distinctions. After this exchange the DMC must move swiftly to make their final decision on termination or to arrange another decision point in the near future. The

future decision point should only come about because of further information to be obtained by the sponsor that the DMC considers relevant.

Should the DMC make a recommendation for termination, the sponsor will decide whether to accept the recommendation. Trial termination will follow immediately in most cases. If the trial is to be terminated, a decision must be made about possible continuation of patients currently on the experimental arm who are tolerating the drug and receiving benefit. The sponsor may either terminate the trial for all patients or leave the decision of continuation up to each investigator and his or her patient to decide. Any patient who continues will have to be reconsented.

When an important decision such as trial termination is being considered, it is important that the DMC report to the same person(s) they normally report to after a DMC meeting. If they are to report to a senior officer of the sponsor if termination is recommended and a clinical research director, they should do so; otherwise an awkward situation occurs when the DMC chair tells the clinical research director that he or she cannot tell this person the outcome of the meeting but rather must contact the senior officer. Rumors will spread through the company very quickly in this case. If a senior officer wants to be first to know of trial termination or other important decision, then he or she must be the person to contact after routine meetings as well.

7.4.6 Unmasking the Sponsor

All of the responses to SAEs of concern described above have the potential of unmasking the sponsor. If the DMC is itself masked, the sponsor would not necessarily be unmasked in being asked to write a "Dear Investigator" letter, modify informed consent, and so on. If it is known that the DMC is unmasked, this would be a signal to the sponsor that most patients with a particular SAE are on the experimental treatment arm. However, this is not a major problem because this information is limited to a few patients and the sponsor would have these clues anyway in observing SAEs expected to occur on the experimental arm. Prior to contacting the sponsor the DMC must decide on a wording of the recommendation that will minimize unmasking the sponsor. It is usually the DMC chair who would have the responsibility of contacting the sponsor and providing the recommendation. Upon hearing the decision the sponsor will usually assume that the DMC has observed a treatment difference that has motivated this change. The DMC should not reveal whether they have unmasked and should not provide more information other than they think the indicated steps are in order. Statements such as "the DMC feels that a 'Dear Investigator' letter is in order" or "the DMC is recommending a protocol modification" are sufficient.

7.5 Information beyond the Present Trial

The DMC may utilize safety evidence from other trials as part of their decision-making process. For data from the same drug the sources of information would be in the investigator brochure issued by the sponsor. However, more detail may be required than what initially appears in this document.

The question arises if the DMC should seek data from other drugs and other indications. It is not the DMC's job to review data from other trials, but when important safety decisions must be made, it is not the DMC's job to ignore such data either. A committee concerned with serious adverse events on a trial of rofecoxib (Baron, Sandler, Bresalier et al., 2006) might choose to seek data on the related drug celecoxib (Lee, Ji, and Song, 2007). Similarly, a revelation of a postmarket safety issue for diabetes drugs rosiglitazone and pioglitazone (Devchand, 2008) might trigger some extra vigilance by a DMC for a clinical trial for muraglitazar (Nissen, Wolski and Topol, 2005) for the same indication. The committee may seek data from other indications of the same drug. A DMC meeting on rituximab for rheumatoid arthritis (Cohen, Emery, Greenwald et al., 2006) may review data previously collected on rituximab in non-Hodgkin's lymphoma (van Oers, Klasa, Marcus et al., 2006) beyond what might be in the investigator brochure. Although the open label extension studies are not controlled trials, DMC members might want to search these studies for further consideration of AEs that arise in the controlled confirmatory trial. In seeking data from other trials, the DMC must seek trials that are contemporaneous with their trial. As Ioannidis, Mulrow and Goodman (2006) point out more recent trials may use new technologies to assess adverse events such as endoscopy for gastrointestinal bleeding versus the previous method of clinical evaluation. They also caution that trials of drugs used in the past may have had low discontinuation rates due to AEs because there were no or fewer alternative treatments at that time. The current trial may have a higher discontinuation rate because of the availability of more alternative treatments. The older trial might have higher AE rates than the current trial due to longer exposure because fewer alternative treatments were available at the time. Also, differentials in frequency and timing of follow-up visits might affect between-trial differences in level of adverse events reported. Table 7.3 presents a summary of cautions in using data from previous trials in DMC decisions. For clinical trials that enroll normal healthy volunteers, such as infectious disease vaccine trials, some DMCs may consider comparing SAEs such as trial death rates with those from vital statistics. Here again there are cautions about comparability. Clinical trial populations are expected to be healthier than the general population because these subjects are healthy enough

TABLE 7.3: Cautions about Using Data from Previous Clinical Trials of Same or Similar Drugs to Compare with Data in Current Trial

1. Previous trials may have had different eligibility requirements and may have allowed different concomitant medications.
2. Current trial may use newer techniques for finding AEs (e.g., endoscopy, CT scans) or newer adverse event dictionaries (e.g., MedDRA).
3. Drugs used in past may have low discontinuation rates due to fewer alternative treatments than present drug or may have higher AE rates due to longer exposure because of fewer alternatives.
4. There may have been differences in frequency and timing of follow-up visits between trials.
5. There may have been differences in use of spontaneous or solicited adverse event reporting.

to make the regular visits required by the trial and because, as trial participants, they are being examined and followed by clinicians more frequently than those of the same age–gender group in the general population.

These steps would be done only when an important safety decision must be made and where members felt that an extra comfort level from additional data was necessary before a decision could be made.

7.6 Meta-Analysis

Closely related to the issue of utilizing data outside the trial is the use of the statistical method of meta-analysis (Whitehead, 2002). The classic definition of meta-analysis is "the statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings" (Glass, 1976). In the clinical trials context meta-analysis entails obtaining either data from published trials or the raw data from previous trials. Then sophisticated statistical methods are used to integrate the data across trials to get combined estimates of efficacy or safety.

A DMC might be presented a meta-analysis by a sponsor or request that one be done to augment a safety analysis. Reasons for this could be that safety data exist on other trials within this clinical program or elsewhere for other indications of the experimental drug. Besides a more precise estimate of AE incidence meta-analysis might permit subgroup analyses that are not possible with the data from the current trial. The sponsor, or preferably an independent group, might perform a meta-analysis when a serious safety concern arises indicating that this occurrence is a chance outlier and a meta- analysis would provide a more realistic estimate. DMCs might see meta-analyses produced by the sponsor if they are reviewing the integrated summary of safety. Here the sponsor might use meta-analysis to refine the dose-response relationship or to see how an overall safety effect might hold among subgroups of patients.

The first consideration in reviewing a meta-analysis is whether it is *retrospective* or *prospective.* A retrospective meta-analysis is performed on data extracted from the literature. A prospective meta-analysis is performed on the actual raw data from the various studies. Patient level data will be available in a prospective analysis but not on an individual basis in retrospective analyses.

Retrospective meta-analysis is highly vulnerable to the publication bias described above. The results of negative trials might never be published, and those trials might have had an unfavorable safety profile for the experimental drug. The data needed for analysis may not be available from all trials or may not be in the same format (e.g., adverse events might be defined differently, age groups might be presented in different intervals). The definitions needed on adverse event classification and severity might not be included in the publications.

Many of these issues exist for prospective meta-analysis as well. The issues in using data from outside the trial summarized in the previous section and in Table 7.3 apply to meta-analysis as well. DMC members should ask the sponsor if they prepared a meta-analysis protocol before embarking on the analysis, where the preanalysis procedures for publication selection, trial qualification, and so on, should be described.

Berlin (2008) discusses several additional issues in the use of meta-analysis in safety assessment. He is concerned that meta-analyses might confound dose and indication. An antiepileptic drug might be used at a lower dose for migraines than for epilepsy. Combining these data might lead to misleading results. If dose differs by gender, researchers must make sure that dose-gender data for the required AEs are available in the literature or in all prospective trials under consideration. It would be important to determine if any of the trials being combined have an active control and if the active control is the same across these trials. A case has been made previously for reporting patient-years of follow-up in all trials. Will these data be available in all trials under consideration?

Meta-analysis can be done with the treatment difference being estimated considered *fixed* or varying in a *random* manner across trials (DerSimonian and Laird, 1986). A test of heterogeneity across trials would be useful in making this decision, and the sponsor should provide these results. Random effects models introduce an extra source of variation which might make it more difficult to find a statistically significant treatment difference in AE incidence. The choice between fixed and random should be justified by the sponsor. Tests of sensitivity to assumptions are helpful to see the extent that conclusions are influenced by classifications, trial inclusions/exclusions, fixed/random, and so on.

Table 7.4 presents a checklist of items DMC members might consider in reviewing meta-analyses of safety data across clinical trials.

TABLE 7.4: Checklist for the Review of Meta-Analyses of Safety Data across Clinical Trials

1. What are the stated objectives of the meta-analysis?
2. Was the meta-analysis prepared by sponsor staff or an independent organization?
3. Was a preanalysis meta-analysis protocol prepared stating how publications and trials were selected, methods of analysis, etc.
4. For retrospective meta-analysis:
   a. How were publications selected?
   b. Is there reason to suspect publication bias?
   c. Can the extent of bias be estimated in some way?
5. All issues in Table 7.3 above apply to review of meta-analyses
6. By what criteria did trials qualify for inclusion?
7. What trials were considered but excluded from analysis? For what reasons?
8. How are adverse events classified and graded across trials?
9. Are doses and indications being combined?
10. Do some of the trials use active controls? Are they the same active controls across trials?
11. Was a test of heterogeneity performed? What are the results?
12. Which factors are considered fixed and which are considered random? What is the justification for this decision?
13. Were tests of sensitivity done to see if there is consistency in conclusions across different definitions of inclusion, fixed/random, etc.?

7.7 Final Meeting

Under the assumption that the trial was not terminated due to safety, we come to DMC decisions at or shortly after the final DMC meeting. The DMC's precise role at this point will be defined by the DMC charter. The following are described under the assumption that these tasks have been previously granted to the DMC.

In some cases the DMC will have the responsibility to review the safety data in the submissions to the regulatory agency. Much of this would be review of tables similar to those already reviewed by the DMC. However, the integrated summary of safety will also be included and DMC review is advisable. This section of the submission will summarize experience of the drug over all studies (Fairweather, 1996; Weihrauch and Kubler, 2002). There should be no surprises here either. The submission will have a description of the DMC, membership, meeting frequency, and so on. The DMC chair can check that this information is correct. Closely related would be a review of safety slides to be presented at the advisory committee meeting. This meeting will take place after submission of the new drug application but before the regulatory agency must make its decision on approval.

Some DMCs review the proposed wording of safety in the package insert. The regulatory agency staff will attempt to ensure that the important safety information is presented in a form that will be useful to practicing physicians, but DMC members may want to comment.

Manuscripts on trial results will be prepared, and if specified in the DMC charter, it will be important that the DMC members review these manuscripts prior to publication. First it will be important to see that there will be a published paper even when the efficacy results are negative. The issue of publication bias is well known (Dickersin, Olson, Rennie et al., 2002; Dickersin and Rennie, 2003). Ioannidis and Lau (2002) have written on shortcomings of many clinical trial papers in the literature with regard to reporting of safety results. They are especially critical of papers that do not report safety data at all, minimize safety data when there are positive efficacy results or report safety as only "medication was well-tolerated." Adverse events are frequently not reported for indications such as the common cold or dry eye because they are not serious and may be mild in nature. This information is useful to clinicians and patients with these indications as the SAEs are to oncology and cardiovascular patients. Ioannidis, Evans, Getzsche et al. (2004) present further recommendations on reporting safety results. A checklist of safety items to look for in a manuscript is presented in Table 7.5.

7.8 Special Problems with Infant Pharma Companies

DMCs must do their job regardless of the size of the trial sponsor. Infant Pharma companies will be especially concerned about the financial implications of trial termination or even protocol modification or "Dear Investigator" letters. When these companies are public and have only one product in development, any action taken

TABLE 7.5: Checklist for Reporting Safety Results in Clinical Trial Manuscripts

1. Mention of safety results in abstract.
2. Were investigators who assessed AEs masked to treatment assignment?
3. Withdrawals due to toxicity and types of AEs that caused withdrawals.
4. Incidence of laboratory-determined toxicity.
5. Safety tables of incidence of all SAEs and those AEs of interest to the community by severity and treatment group, not just the most frequently occurring AEs.
6. Patient years of follow-up should be reported for each treatment group.
7. Assay sensitivity—what odds ratios or hazard ratios can be detected by sample size?
8. Severity should be reported as well as the severity scale used.
9. Table of rare and unexpected AEs regardless of severity.
10. Was AE collection spontaneous or solicited? What was frequency of data collection?
11. Avoid broad categories that do not permit clinical interpretation such as "gastrointestinal AEs," "skin disorders."

by a DMC may be considered material and thus must be reported in a press release as required in the United States by the Securities and Exchange Commission. When companies are private, the sponsor may still be required by their Board of Directors to report this information to investors. DMCs should not be intimidated by this requirement. The requirement is to protect investors and the DMC exists to protect patient safety. However, the DMC should be aware of this risk and avoid hasty decisions when serving an Infant Pharma company.

7.9 Conclusion

We have now followed the life cycle of a DMC from concept to creation to final meeting. We have learned a lot about clinical, statistical, and bias issues along the way. However, the pharmaceutical clinical trial arena is constantly changing with new regulations, clinical advances, new statistical methodology, and so on. We will try to apply what we have already learned to some of these emerging issues in the next chapter.

DMCounselor

Q7.1 I am sitting on a DMC where the accrual is very slow and is not likely to ever reach the sample size to establish efficacy that appears in the protocol. Can the DMC terminate the trial on the basis that patients are being exposed to a drug where we are not likely to ever learn about its efficacy?

A The answer to this question is yes, but the DMC must first work with the sponsor on analyzing the reasons for poor accrual, suggest new investigators, change in eligibility requirements, and so on. The DMC and the sponsor should come to an agreement of how long to wait for accrual to improve before the trial is terminated. The more severe and serious the AEs reported, the shorter should be the waiting time. While waiting for improvement you might ask the DAC statistician to make a conditional power or predictive power calculation. This might guide a decision later, but this should not be thought of as a strictly statistical issue.

Q7.2 I am serving on a DMC where the active control was approved only a year ago. With unmasking we are seeing an excess of SAEs in the control group. Can we terminate the trial because of safety issues in the control group?

A The answer is yes. The DMC exists to protect patient safety regardless of what treatment arm they were assigned. The DMC may not have to terminate the trial. The members can work with the sponsor to choose another active control, and if an ethical alternative exists, this should be done as soon as possible. Many SAEs do not present themselves until postmarket so this situation is not unusual. The DMC should make sure that these control group SAEs are reported to *MedWatch* or other appropriate regulatory postmarket surveillance system (MedWatch homepage, 2007).

Q7.3 I am a pediatric hematologist serving on a DMC for an experimental treatment for pediatric epilepsy. I have been placed on this DMC due to expected hematologic AEs. I have chaired several DMCs for other sponsors for treatments in pediatric anemia. The chair of this DMC is a pediatric epileptologist. I have noticed that he has done quite a bit of work for this sponsor over the years and is always having small talk with sponsor representatives about other sponsor personnel that he knows before our open session begins. The other MD members of the DMC are all pediatric epileptologists and have known one another for many years. I am the new kid on the block. We are masked to treatment but during our closed session the chair often plays down the importance of SAEs that we feel might be on the experimental arm or he attributes them to the active control. The other pediatric neurologists generally agree with him instantaneously. I have asked them to look closer at some of these SAEs but they are not hematological in nature so they feel that they know more about this than I do. We seem to be in a decision-making deadlock. What can I do about this?

A First you and the biostatistician member of your DMC should remind the other members that you are free to comment on any safety issue that you think is important regardless of board certification. Ask the other members why they think the SAEs are due to the active control. Try to show any logical failures in their analysis if you can. Consult the literature and the active control package insert to learn of the SAEs expected. You can also consult with a pediatric epileptologist at your institution to see what he or she thinks of these SAEs and the active control. Of course you cannot tell this person why you are asking but that should not be necessary. You can ask to be unmasked to the treatment assignments of just the patients that have the SAEs of your concern. Typically the whole DMC should be unmasked. If they refuse, at least you can be unmasked. If the SAEs are occurring more frequently on the experimental arm, these steps should bring them out.

Q7.4 I am a sponsor representative working with a DMC in infectious disease. Our DMC has asked the DAC to prepare several ad hoc safety tables to address what they describe as an important concern. I approved the generation of these tables, but three weeks have passed since the tables were completed and the DMC members can't agree on a date for their

teleconference to go over the reports. We at the sponsor are concerned that patients are at risk and rumors are flying around the office which means that this information might have also reached the financial community. Part of the problem may be that because so many countries are represented in this trial, we have six physicians on the committee, each representing a country being served.

A I would not recommend having fewer members on your DMC just because of problems in having ad hoc meetings. It is probably important to have six physician members but you must have 6 committed members who are contactable wherever in the world they may be at the time an issue arises. However, the easiest way to address this problem might be to schedule a teleconference day before you approve the extra work for the DAC. The DAC must first tell you how much time they need to prepare the requested tables, listings, graphs, and so on. Once you have a commitment from all members to meet shortly after delivery, you can approve the work order.

Q7.5 I am chair of a DMC working on a cardiovascular indication. We have asked the DAC for several ad hoc reports to aid in interpretation of safety concerns. The DAC tells us that they need several weeks for each table because of the need for validation of the software and cleaning of the data. We don't feel that we can take that much time given that patients are at risk. Can we overrule the need for validation and data quality control?

A You would not want to eliminate all software validation and data cleaning prior to generating the tables you need. You would not want to make decisions based on faulty data. In Table 3.1 we indicated that the extent of software validation and data quality control prior to data review was something to be entered into the DMC charter before the trial begins. If the DAC has SOPs calling for certain validation procedures, it may be impossible to modify them. If they do not have SOPs for software validation for DMC operations, they should not have been selected as a DAC in the first place. The best procedure is to discuss a reasonable validation approach that gives the quality the DMC needs for their deliberations in the time frame that seems appropriate. If SOPs are being violated, a note can be written to the file detailing the reason for departure and the procedures that were followed.

Q7.6 I am an ophthalmologist serving on a DMC for an infectious disease of the eye. The other members are a retinologist who is serving as chair, an internist, and a biostatistician. We saw evidence of a very serious adverse event six weeks ago and I feel we should have told the sponsor to stop the trial then. The other members are asking for more information on the SAE reports, reading the literature including papers on animal studies that I do not feel them qualified to evaluate. Meanwhile patients are at risk. I am the only member of this DMC who treats patients of this type every day. I am growing frustrated at the amount of time the others have taken to investigate things that I know are irrelevant. Should I just contact the sponsor myself?

A Committee work has never been the most efficient way of doing business but you should not undervalue what the other members are doing just because it takes time. They may have a different, but nevertheless, important perspective. If the trial were to stop now, it will be difficult to start it again should your fellow committee members come up with information that might explain the SAEs and lead to recommendations for preventing these events. On the other hand there may be people more qualified with more resources within the sponsor to deal with this issue. You ought to work with your fellow DMC members to establish a deadline for their research, and then the DMC should brief the pharmacovigilence unit at the sponsor. The members of this group can be unmasked

and can look into this matter without informing the study team of what is going on until a trial decision is needed. Also, the DMC does not have to tell this group that the trial should stop but merely that the sponsor needs to give the DMC reason why the trial should continue. You will find that the pharmacovigilence group has experience in dealing with these issues and should know how to proceed.

Q7.7 I am a cardiologist from Denmark serving as chair of a DMC for a cardiovascular drug. The trial is being conducted in Europe. The sponsor is collecting cost and quality of life data during this confirmatory trial in order to present to the UK NICE organization and the German IQWiG. The problem is that the sponsor is asking us to advise them on the quality and treatment differences of the cost and quality of life data along with the safety data. The cost and quality of life data take up so much of our time that I feel that we don't devote enough time to the safety data, which we thought was our highest priority. In addition the North American members of our DMC are not helping with the cost and quality of life data effort. They consider the data to be inadequate to answer any questions on cost effectiveness. How can I get my committee back on track?

A You have encountered some of the problems of intercontinental differences in the drug approval process. Some European countries have the additional step of cost effectiveness and North Americans are not used to cost-effectiveness being measured in the manner you describe. I do feel that safety should be a priority and I am sure that sponsor would agree. The sponsor should have made the cost effectiveness responsibilities clear at the outset. At this point you, as chair, should discuss with the sponsor representative the possibility of their assembling another committee to review the cost and quality of life data. A one-person committee consisting of a knowledgeable consultant might be adequate for this purpose. If so, this individual could be added to the DMC to do this review while the rest of you work on safety.

Q7.8 I am serving as chair of a DMC for colorectal cancer. Since our last meeting, the predetermined boundary for indication of futility was reached, so the DAC presented a very minimal safety analysis under the assumption that we would terminate the trial due to futility. Under what authority can they do that?

A This was bad practice on the part of the DAC. If the DAC did not receive instructions from the DMC regarding a revised analysis, they should have provided the analysis that had been prespecified. Also, regardless of the crossing of efficacy boundaries, there are still safety concerns to be discussed. If the trial is to be terminated, the DMC must consider continual treatment of patients in the experimental group and, despite futility, whether to recommend crossover to patients in the control group. The safety tables you prespecified would be an important resource for those recommendations. You must demand that the usual tables be sent to the DMC as soon as possible.

Q7.9 I am a neurologist serving on a DMC for a Parkinson's disease clinical trial. At our last meeting we passed the futility boundary, but our chair wanted to overrule the boundary because she felt that the experimental drug had a favorable safety profile and the trial would still serve as a noninferiority trial. Our biostatistical member felt this was OK. This did not sound right to me, so I asked a statistical colleague at my institution who consults often with pharmaceutical firms, and he said that arbitrary unplanned switching from superiority to noninferiority could not be done. For confidentiality reasons I did not reveal the reason for my question to this colleague. I could call this to the attention of our biostatistical member, but I could not argue the matter with her, and due to confidentiality, I cannot put her in touch with my academic colleague. Where do I go from here?

A There are a number of issues here. First, it is not clear that your DMC has the right to overrule a futility boundary and proceed with a plan B. Your committee's task is to inform the sponsor that the futility boundary has been reached. You can recommend continuation because of the favorable safety profile but surely someone on the sponsor side would know, as your academic colleague correctly stated, that switching from superiority to noninferiority can be done only by an approved adaptive design and with a preapproved indifference margin. The sponsor of your trial would be better served to terminate this superiority trial that your DMC is working on and design a separate noninferiority trial if that makes clinical and market sense to them. Your DMC's biostatistical member appears to be uninformed about certain statistical/regulatory technical details common to the pharmaceutical industry. This is not uncommon when academic biostatisticians with little previous connection to the pharmaceutical industry are placed on DMCs. After your trial is terminated, you and your academic colleague should inform her. This is all part of the training needed for the success of future DMCs.

Chapter 8

Emerging Issues

Bullets to Remember

- Adaptive designs are heavily dependent on efficacy data. This presents a problem for DMCs when efficacy and safety data are divergent.
- In noninferiority trials any safety differential that favors the active control may cause the DMC to declare the experimental drug inferior to the active control and want to terminate the trial regardless of efficacy data.
- Skin rash is an adverse event but is also a biomarker for efficacy for certain oncology drugs. Thus, sponsor staff processing routine AE reports can be unmasked.
- The time has come to at least discuss formal training and certification for DMC members.
- Sponsor and DMC chair should reach some agreement on cost control for DMC operations and requests before the trial begins.
- DMCs have the responsibility to make sure there is continued independent review of patient safety as drugs are licensed to other companies or during mergers and acquisitions.
- Apart from newly developed conflict of interest resignation from a DMC should be a last resort.

8.1 Introduction

This chapter investigates recent trends in the pharmaceutical industry and how they might affect the operations of a DMC. The trends are divided into two broad categories—those related to changes in technology or company organization and those related to the maturity of DMCs as a clinical trial component.

8.2 Issues in Technology

8.2.1 Adaptive Designs

The pharmaceutical industry has realized that there is a need to accelerate drug development because clinical trial design, analysis, and the regulatory process have not kept up with advances in technology. There are more compounds under investigation for more indications than has ever been the case previously. Many treatments will not make it to approval, and sponsors want to eliminate these candidates as soon as possible in the development cycle. Those treatments that will be approved will have relatively short lifetimes due to competition and rapid technological advance. In the United States, the FDA has acknowledged this need with their Critical Path Initiative (U.S. FDA, 2004). In this program, the FDA is encouraging research in biomarkers, genetics, and statistical/clinical trial methodology to help in the needed acceleration. One of the principal statistical ideas for accelerating the clinical trial process is adaptive designs. Gallo and Krams (2006) and Dragalin (2006) have provided a good overview of the types of adaptive designs under consideration. An adaptive design is one that uses data accumulating during the clinical trial to change sample size, alter hypothesis, drop treatment groups, make a "seamless" transition from a phase II to a phase III trial, and so on while preserving statistical properties such as Type I and Type II error. Obviously drug development efficiency is gained in adapting a trial rather than beginning an additional trial using information from the previous trial. Discussion of the pros and cons of these proposed methods is beyond the scope of this book. However, Herson (2008) has discussed the effects of various types of adaptive designs on DMCs involved principally in safety monitoring. Gallo (2006) has offered some ideas for alterations in trial monitoring that might be necessary for adaptive designs.

A problem arises because adaptation is based on efficacy alone and the DMC is involved in periodic safety assessments independent of efficacy. To avoid unmasking and bias, the sponsor is unaware of adaptive changes being made. The challenge is what to do if safety concerns are in conflict with adaptive changes. Some examples follow.

8.2.1.1 Dropping a Dose or Treatment Group

Suppose a clinical trial with three experimental treatments or doses of the same treatment are being compared to a control group in an adaptive design trial that permits dropping an experimental group on the basis of efficacy. Clearly if the DMC is given word that a group is being dropped, they immediately know that this group was not the control group. However, we have already stated that there is no major problem in the DMC's being unmasked to treatment. The problem occurs if the DMC has developed safety concerns about the groups that will remain in the trial but was comfortable with the developing safety profile of the treatment group being dropped. Upon being informed by the DAC biostatistician that the conditions for dropping this "safety-friendly" group have been met, the DMC might suddenly be discussing terminating the entire trial rather than continue with treatments that may demonstrate efficacy but do not have a favorable safety profile. This being an interim

analysis, it may be too early for the DMC to look at efficacy data and evaluate risk versus benefit for the remaining groups. However, they must be careful in discussing a decision with the sponsor because the information divulged can introduce bias if the trial is to continue. The best solution would be for the DMC to begin a dialogue with the sponsor about terminating the trial and not make a recommendation until sufficient discussion has taken place. This discussion should take place with sponsor staff not directly involved in the trial, and only limited information about treatment performance should be given. A discussion with noninvolved sponsor staff might be possible for Big or Middle Pharma but not for Infant Pharma where everybody in the company has the potential of an emotional tie with the outcome of the trial. In these cases CRO staff or the staff of a Big Pharma corporate partner may be able to participate in this dialogue.

8.2.1.2 Adaptive Assignment to Treatment Group

Under this scheme the random allocation to treatment groups changes dynamically according to accumulating efficacy data, with the treatments showing more efficacy potential being allocated more patients. Here the DMC will be very aware of efficacy because the safety tables they review will have more patients on the more favorable arms. Concerns may arise if these so-called favorable arms are also considered most toxic. The imbalance created by adaptive assignment may not permit adequate comparison of treatment groups for risk versus benefit. One solution would be for the DMC to recommend suspension of allocation change when imbalance becomes too large. Allocation can be changed again after the safety profile is considered better understood by the DMC.

For the remaining adaptive design sections, let us assume that there are only two treatment groups—experimental and control.

8.2.1.3 Changing Objectives: Superiority to Noninferiority

Much has been written about noninferiority trials and the statistical methodology involved (D'Agostino, Massaro and Sullivan, 2003). In the noninferiority trial the objective is to show that the experimental treatment is at least not much worse in efficacy than an already approved drug for this indication (active control). Presumably there is some benefit to the experimental treatment in terms of safety, convenience, cost, etc. In a superiority trial in oncology, DMC physicians might be willing to accept grade 3–4 nausea and vomiting hoping that an improvement in overall survival will be demonstrated at the end of the trial. With a switch in objectives to noninferiority the DMC will be told that the sponsor does not expect to demonstrate efficacy superiority against the active control. The physician DMC member mindset will often change at this point. Now grade 3–4 nausea and vomiting in one arm (and known not to exist in the active control arm) becomes a reason to deem these treatments as not equivalent regardless of efficacy data. Some physician members will claim that inferiority of the experimental treatment has been demonstrated (not superior in efficacy, more toxic than control), and hence, there is no reason to continue the trial. This situation would be less likely to happen if an adaptive change objectives design was not being used because in those cases the sponsor would usually have sufficient evidence of a

favorable safety profile before beginning the trial. In an adaptive design the safety profile is usually not that well known at the time the trial begins, and hence, this kind of situation can occur occasionally. It is not clear what the communication between the DMC and sponsor should be in this case. The DMC may not have a safety *concern* but more of an observation of a safety *differential* between treatments which would deem the experimental treatment inferior to the active control. A DMC recommendation of trial termination appears appropriate in this case.

8.2.1.4 Seamless Transition: Phase II to Phase III

Seamless transition can take place where efficacy data support proceeding from phase II to phase III. The patients already enrolled in the phase II trial become part of the phase III trial, and the sample size is increased to pivotal trial levels. At the time of transition the DMC may have already developed safety concerns and believe that it is inappropriate to increase sample size. This situation appears simpler than those cited above and is one that occurs in the practice of sample size reestimation (Chuang-Stein, Anderson, Gallo et al., 2006). In most protocols that allow for sample size reestimation through an interim analysis, it is understood that sample size increase can only take place if the sponsor has no safety concerns that would preclude increasing sample size. Thus seamless transition can easily be absorbed by DMC operations under sample size reestimation.

8.2.1.5 Change in Effect Size of Interest

A popular method of sample size reestimation is the method of Cui, Hung, and Wang (1999). This scheme uses a conditional power calculation at a planned interim analysis to compute the sample size necessary to deliver the required power for the *observed* effect size. An odds ratio of 2.0 in favor of the experimental treatment may have been agreed upon at the start of the trial, but an odds ratio of 1.4 is observed at the interim analysis. An increase in sample size at this point is indicative of a change to a smaller effect size than that agreed upon as clinically significant at the beginning of the trial. The sponsor realizes that if this smaller effect size is found statistically significant at the end of the trial, they will have to convince regulatory authorities that the revised effect size is of clinical significance. However, DMC members will be aware of the change in effect size at the interim analysis. It has already been mentioned that risk versus benefit is a difficult task during or even at the conclusion of a single pivotal trial. It is even more difficult when the benefit (effect size of interest) is a moving target. Some DMC members may feel that the emerging safety profile cannot be justified with the smaller effect size. This would be reason to terminate the trial. Under these circumstances the DMC may recommend to the sponsor that safety considerations dictate that the effect size of interest not be changed. There would not be a reason to indicate what the safety concerns are at this point. The sponsor may choose to terminate the trial due to the futility of finding the original effect size statistically significant at the end of the trial.

A summary of DMC communications and actions in adaptive designs is found in Table 8.1.

TABLE 8.1: Summary of DMC Issues and Actions in Trials with Efficacy-Based Adaptation

Efficacy-Based Adaptation Type[a]	Issue	Data Monitoring Committee Action
Dropping dose or treatment group	The dose/treatment group being dropped might be the only one in the trial with an acceptable safety profile	Begin a dialogue with sponsor representative on need for trial termination
Adaptive assignment to treatment group	Random treatment allocation is favoring a treatment arm with an unfavorable safety profile	Either ask sponsor to change allocation to equal per treatment group until safety profile becomes clear, recommend elimination of unsafe treatment group or terminate trial
Changing objectives—superiority to noninferiority	Regardless of efficacy unfavorable safety profile in experimental group means treatments are not equivalent	Recommend trial termination
Seamless transition—phase II to phase III	Sponsor is aware of this transition and will consult DMC for possible safety concern before proceeding to phase III.	Recommend transition if no safety concern
Changing effect size of interest	DMC members may feel that the emerging safety profile of the experimental treatment cannot be justified by a decreased efficacy effect size of interest derived by a conditional power calculation at an interim analysis.	Recommend that the effect size of interest remain as originally planned if that effect size can be justified by safety risk or terminate the trial.

[a] Communicated to Data Monitoring Committee by Data Analysis Center staff.

Source: Herson, J. (2008) Coordinating data monitoring committees and adaptive clinical trial designs, *Drug Information Journal,* 42, 297–301.

8.2.1.6 Further Thoughts on Adaptive Designs

It must be clear that adaptive clinical trial designs, other than sample size reestimation, are a new concept for pivotal trials. There are many details to be worked out to satisfy concerns of sponsors, investigators, and regulators. The discussion above is only

intended to mention some concerns that a DMC focusing on safety issues might face when adaptation based on efficacy is about to take place. Besides concern when the experimental treatment's efficacy and safety characteristics appear to diverge, there is the problem of determining proper DMC-sponsor communication about the concerns that does not introduce bias into trial operations. It probably makes sense for protocols to provide for a period of "white space" (period of time reserved for planning) whenever adaptation takes place. During this period the DMC would be asked to comment on safety concerns before the adaptation could take place. The problem remains of what sponsor unit would receive this report. Later in this chapter we will discuss the appearance of internal safety review committees which are beginning to appear in Big Pharma. These committees consist of physicians and biostatisticians employed by the company but not working on this trial. The rationale for this will be discussed in the following paragraphs. For Big Pharma using the internal committee, the DMC could presumably report to this body. Such a unit would not exist in Infant Pharma, and for these sponsors, the DMC will have to issue a carefully worded statement. One alternative is for adaptive design trials not to be used at all until previous trials establish the safety profile. For sponsors of all sizes physicians may be reluctant to become involved in DMCs for pivotal trials when the investigator brochure for the experimental drug is sparse on safety.

The biostatistician member and the DAC biostatistician must both believe in and be well versed in adaptive methods and be able to explain them to the physician members. The physician members must also be comfortable with the adaptive approach before agreeing to serve on the DMC. Before any data review meeting where adaptation is a possibility, the two biostatisticians must lead the physician members through a decision matrix of steps that will take place, contingent on the data to be reviewed. The rationale for each branch of the matrix must be carefully explained. This is best done before members receive the data to be reviewed at the meeting. A conference call can be used for this purpose and timed to take place before data are sent to members.

8.2.2 Real-Time SAE Reporting via the Internet

Many sponsors are using Internet-based collaboration software for DMC support. This software allows DMC members to view various documents throughout the trial. In some clinical trials individual SAE reports reach DMC members via an e-mail request to go to a Web site to view the report. This e-mail is generated as soon as the SAE report arrives at the sponsor. This type of real-time reporting is useful after a DMC has developed a safety concern such as early deaths or cardiovascular disease, and decides to monitor the situation closely before making a final decision. However, reporting all SAE reports in this way has the potential effect of creating attitudes among DMC members before reviewing cohort data. These attitudes could result in premature unmasking if the DMC is to remain masked or becoming overly conscious of a certain SAE type and ignoring others. A compromise might be for the real-time SAE reports to go only to the DMC chair or another mutually agreed upon member. The chair would decide if there is reason for an ad hoc meeting or for additional data review tables for the next DMC meeting that might clarify the situation.

8.2.3 Causal Inference

An exciting area of statistical research is the development of methods of causal inference (Frangakis and Rubin, 2002; Rosenbaum and Rubin, 1983). These methods use the statistical techniques of propensity scores and principal stratification to allow for adjustment of post-treatment variables that allow for conclusions of causation. One safety application that a DMC might encounter would be a post hoc analysis of safety based on variables measured during the trial. Suppose a DMC is monitoring a clinical trial for an oral pain medication and discovers a rate of cardiovascular disease in the experimental group much higher than that in the placebo group. It is also discovered that some patients overdosed the medication. Sponsor staff might perform a causal analysis of the effect of overdosing on the outcome of cardiovascular disease even though patients were not randomized on the basis of compliance with dose. This could lead to protracted discussions if the causal analysis revealed that it was overdosing and not the experimental treatment that caused the cardiovascular disease. Presumably there would still be evidence against the experimental treatment such as overdosing being caused by perceived lack of efficacy by the patient. In any case the DMC biostatistician might not be selected on the basis of familiarity with causal inference methodology because, at the beginning of the trial, nobody thought such an analysis would be performed. This could be especially problematic if this analysis were done in response to a DMC observation of excess cardiovascular risk at an interim analysis rather than at the end of the trial. The resolution of this issue will depend on the specifics of the drug and the indication. However, if such an analysis were performed at the end of the trial, the DMC would, presumably, have no problem with the analysis appearing in a manuscript for the trial, and the regulatory agency would have input into how this analysis might be presented in a package insert. If this analysis occurs at interim analysis, the DMC would have to take it into consideration in a decision on trial termination.

8.2.4 Biomarkers

The discovery of biomarkers for patient response and their use as surrogate endpoints in clinical trials have the potential of increasing efficiency in clinical trials. However, recent clinical trials in oncology have shown that the epidermal growth factor drugs cetuximab (Lenz, Van Cutsem, Khambata-Ford et al., 2006) and erlotinib (Wacker, Nagrani, Weinberg et al., 2007) are associated with skin rashes and the severity of the rash is correlated with survival time. Thus observation of an adverse event reveals information on efficacy. This is a problem for sponsor pharmacovigilence staffs. They are supposed to review adverse events but be masked to efficacy. There is no major problem of DMC members' being aware of these data but they must be careful about discussing concerns about skin rashes with sponsor staff. The pharmacovigilence unit is a vital part of the safety monitoring process, and they should not be shut out of safety data review. A possible solution is for one pharmacovigilence employee who does not work on oncology to receive all adverse event reports first and remove the skin rash reports. The remaining reports would be given to the oncology pharmacovigilence staff. If the skin rash reviewer sees a problem in volume or severity of skin rashes,

he/she can inform the DMC chair. It is expected that more biomarkers related to efficacy in the form of adverse events will be discovered in the near future and there will be a need for further development of procedures of this type.

8.2.5 Exciting Times Ahead

Technological advances will make it possible for even faster reporting and processing of AEs. There will be a need for simultaneous procedural change to reduce bias in safety monitoring. Research in personalized medicine will eventually call for clinical trials where each patient will have his or her own treatment although coming from a common technology that creates the personalized treatment. This will also call for change in safety monitoring because AEs will occur in small numbers and, perhaps, be more varied than we find with treatments today. DMCs will have to find signals not from frequency but from severity regardless of frequency. Cogent reporting to a sponsor will be a challenge.

8.3 Issues Due to Maturing of DMC Processes and Evolution of the Pharmaceutical Industry

Use of DMCs in safety monitoring has become standard in most pharmaceutical industry confirmatory trials. As a result, questions other than appropriate content of the DMC charter or number of members arise. We will examine some of these issues as well as some that arise from the evolution of the pharmaceutical industry.

8.3.1 Training of DMC Members

In the early days of DMCs in pharmaceutical industry clinical trials and as the DMC concept developed, members received on-the-job training. At first there were few people available to serve on DMCs who had the experience to train others and sponsors were putting together SOPs which, initially, varied considerably from company to company. The guidances from regulatory agencies have helped in harmonization of procedures. The question that arises is whether on-the-job training is still the best preparation for DMC service or whether formal training programs should be organized. The latter would consist of Internet courses, apprenticeships, and certification. Of course those who have served on DMCs would automatically be certified.

Training could take place through courses given online or at meetings of the Drug Information Association, Regulatory Affairs Professionals Association, Association of Clinical Research Professionals, British Institute of Regulatory Affairs, etc. The written syllabus would consist of material presented in this book but role playing and apprenticeship would also be part of an ideal program.

The pros for a formal training program would include that it would allow younger professionals to serve on DMCs who might otherwise not be considered due to lack of

experience. Diversity of DMC members would have many advantages over a smaller aging power elite of DMC members. A database of certified DMC professionals would help sponsors staff DMCs. As North American and European societies become more multicultural so are the patients who volunteer for clinical trials. It is important that this ethnic diversity be represented on DMCs.

The cons for such a program would include potential members' not having the time or being willing to pay the expense for the training out of their own pockets. Apprenticeship training would require that sponsors pay the DMC members-in-training some stipend for meeting attendance and reimbursement for travel expenses. Several people have proposed this to sponsors and been turned down.

DMC training should also be provided to sponsor staff who would be working with DMCs but have no previous experience. Several experienced DMC members have had the experience of doing on-the-job training of sponsor staff who would not know how to conduct open meetings, or what data review tables are required, coordinating data flow between sponsor, DAC and DMC, and so on. It should not be the responsibility of the DMC members to train sponsor staff. Sponsors can create internal programs to train their staff members and apprenticeships consisting of attendance at DMC meetings with more experienced sponsor staff. Other groups who might benefit from training in DMC operations would include persons wishing to perform DMC audits (see section below) and CROs that might want to serve as DACs.

8.3.2 Cost Control

In their deliberations DMCs frequently see the need for ad hoc meetings and ad hoc tables and listings in order to better understand emerging safety issues. They may even ask to have an ad hoc consultant such as a cardiologist or an immunologist/allergist appointed for guidance in interpreting adverse events lying within their expertise. The DMC must be careful in making these requests because they do not want to unmask sponsor staff. As there is more experience with requests of this type sponsors are asking if they should be required to pay for this additional work without having input into the decision. All too often sponsors have been getting bills for thousands of dollars for extra work only to find that the DMC ultimately decided there was no safety concern. At the end of the trial sponsor staff are unmasked and, viewing the data, wonder why the ad hoc requests were made (see *DMCounselor* Q8.2).

Closely related to the cost control issue are ad hoc requests made to the DAC other than just generating additional tables with more granularity in a certain organ classification. For example, there may be requests for conditional power calculations, causal inference methods, or combining MedDRA preferred term codes across SOCs. Besides the cost of these activities, the requests may involve capabilities that the DAC staff does not have because the DAC was selected on the basis of different criteria. This is especially true when the DAC is based at a CRO working under contract with the sponsor. However, it can occur as well when the DAC is a separate unit within the sponsor.

One way of handling this would be for the DMC and the DAC to have a certain discretionary budget for ad hoc requests. This means that the DMC chair would have

to ask the DAC for the cost of the ad hoc requests and determine if the budget would be exceeded. If the budget is to be exceeded, or if consultants must be retained, the DMC will have to make carefully worded requests to the sponsor. The sponsor will have the right to veto the requests, but there may be a middle ground such as reducing costs in other areas, perhaps by eliminating routine tables that are no longer needed. If the sponsor vetoes, their representative should present the sponsor's rationale for the veto to the DMC. The rationale should not be merely that the request is too expensive. Often the sponsor staff will have considerable information on mechanisms of action and pathways that may explain adverse events that DMC members deem important but do not have this deep knowledge of the drug.

8.3.3 DMC Audit

During the process of drug development audits of investigator sites, CROs, sponsor records, manufacturing facilities, and clinical laboratories are common (International Conference on Harmonisation, 1996). The audit process has recently included DMCs. The audits may be performed by the FDA or by the sponsor in anticipation of an FDA audit. Another pharmaceutical firm thinking of licensing the drug the DMC is monitoring might also want to make an audit. A DMC audit would consist of going over the charter and minutes of DMC meetings. The auditor would want to ascertain that the charter was followed, patient safety was under the stewardship of an independent DMC, there were no obvious conflicts of interest among DMC members, and there was no unmasking, accidental or otherwise, that could bias trial conduct. An important document in this audit would be the DMC minutes. These minutes would have been prepared by the DMC secretary or the DAC biostatistician. There would be both open session and closed session minutes. In order for these minutes to be available for inspection at any time there must be a record retention procedure in place. No records should be maintained in the home or office of a DMC member. The open session minutes can be stored by the sponsor clinical staff, but the combination open–closed meeting minutes should be stored at the DAC or at some office of the sponsor separate from those involved in the trial, such as a manufacturing or quality control office.

8.3.4 Internal Safety Review Committee

The internal safety review committee (ISRC) is a fairly new phenomenon in Big Pharma. The ISRC is structured like a DMC with physician members and usually one biostatistician. All members are employees of the sponsor but work in other disease areas. The ISRC follows a charter similar to the DMCs as far as masking is concerned. The DMC meets only periodically and some Big Pharma sponsors feel there is a need to having an internal group meeting more frequently than a DMC can. The ISRC receives the same information as the DMC. The ISRC does not make unilateral decisions. They might hold ad hoc teleconferences with the DMC when important safety issues arise. The biostatistician member of the ISRC is usually the DAC biostatistician for the trial. The physician members of the ISRC might attend closed sessions of the DMC.

The use of an ISRC in a clinical trial brings many changes to the sponsor-DMC relationship and the data flow between sponsor, DAC and DMC. Most importantly it raises the question of who is ultimately responsible for the stewardship of the trial. Having the ISRC as part of DMC discussions on safety might change the direction of decisions because, while some may question whether a DMC is truly independent of the sponsor, surely the ISRC is less independent. This is a new concept and, if an ISRC is to be involved in a trial, the DMC charter should spell out the precise roles of each unit in detail. The DMC should make sure that the regulatory agencies involved are aware of the presence of the ISRC.

The ISRC should be distinguished from pharmacovigilence committees (sometimes called firewall or medical governance committees), which were mentioned earlier. These committees do not regularly review data on the trial as the ISRC does. They exist for consultation with the DMC when serious safety concerns arise which cannot be discussed with the sponsor staff because of unmasking potential.

8.3.5 Mergers and Licensing

During the course of a clinical trial, the sponsor could merge or be acquired by another company or the experimental drug could be licensed to another company. This change in ownership of the product might result in immediate personnel changes in the sponsor, and there may be a period where it is not clear who is in charge at the sponsor or what the plans are for the continuation of the DMC. The new ownership may have different SOPs for a DMC and safety monitoring than the original owner, and the new owner may want to choose their own DMC. In these cases it is the responsibility of the original DMC to see that the trial is not orphaned. The DMC must make sure that independent stewardship of the trial continues. The DMC chair will have the responsibility of finding out who is now in charge and what the plans are for safety monitoring. Unfortunately if the new owner says that the original DMC will not be needed and does not disclose its current plans for safety monitoring, there is not much that the DMC can do to correct this matter. A more typical situation is that there is a new sponsor but the original sponsor must continue the safety monitoring until the trial is completed. This orderly transfer makes a lot of sense and is in the best interest of both companies and the patients. The DMC may find that sponsor staff are less interested in the trial now than they were before the change in ownership, but the DMC chair must not allow this change in attitude to affect DMC attitude or commitment.

8.3.6 Journal Policies Regarding Independent Review

Recently *JAMA* published a policy on publication of results of industry-sponsored clinical trials (Fontanarosa, Flanagin, and DeAngelis, 2005). The new policy stipulates the inclusion of an independent DMC to oversee the trial. This may have come as no surprise. However, the policy statement also included a requirement that an academic-based biostatistician receive the data from the trial, verify the appropriateness of the design and analysis, and plan and make an independent analysis of results. There was no mention of what would happen if the sponsor and independent biostatistician results

did not agree. This is a very controversial policy, and it may never be implemented at least in this form. If there is an independent biostatistical review, DMCs can expect comments on the following types of safety analyses: multiplicity, unplanned analyses, informative censoring and competing risks, assay sensitivity (power), and so on. All of these issues were covered in Chapters 5 and 6 and would not necessarily all wind up in a publication. If comments such as these are generated, DMC members will have to meet with sponsor representatives to decide how to proceed.

8.4 Resignation from a DMC

It is always hoped that a DMC member will not have to resign during the lifetime of the clinical trial. However, there are some obvious reasons for a member having to resign. These would involve conflicts of interest that have arisen from changes that have taken place since the trial began. Examples would be becoming an investigator in a trial for a competing product, becoming an employee of another pharmaceutical firm, becoming more involved with the sponsor as a consultant, and so on. There are personal reasons such as health and increased workload at one's own institution that might also lead to resignation.

DMC members may have serious disagreements with other DMC members or sponsors that cannot be resolved. These disagreements may be over the nature of data analysis, safety concerns, or ethical issues. There may also be personality conflicts between DMC members or with sponsor representatives. Whatever the reason the DMC members must be sure that they have heard the problem of the DMC member's concern and done everything to resolve the problem.

The next step after the resignation would be for the DMC to work with the sponsor to see that a qualified replacement is found, and if possible, the resigning member continues to serve until the replacement is in place.

The DMC charter will typically list criteria by which the sponsor can terminate a DMC member. If this should happen, the procedures of the preceding paragraph would apply.

8.5 Conclusion

We have now gone through a book devoted to the state of the art and best practices in safety monitoring through DMCs in pharmaceutical industry clinical trials. This chapter has brought to our attention that the art and science of safety monitoring is constantly in flux partly because of changes in technology and partly because of the maturing of the DMC process. Drugs that are not approved by the safety monitoring process will go back to the planning stages for some rethinking—perhaps a slightly altered molecule or the same molecule in a different formulation, dose, schedule,

different patient types, different indications, and so on. For those products approved, the safety monitoring process is not over. As we said in Chapter 1 we really start to learn about the safety process and the risk versus benefit tradeoffs of a new drug in the postapproval era. Here community physicians will voluntarily submit adverse event reports to regulatory agencies, and eventually, thought leaders will be able to write and speak about risk versus benefit. However, the history of any new drug begins with the safety profile determined from clinical trials and the DMC plays a vital role in this process.

DMCounselor

Q8.1 I am a biostatistical member of a DMC for a renal cell carcinoma trial. At the first meeting of our DMC I raised questions about an obviously erroneous early termination rule in the protocol and asked that it be revised to a conventional Lan–DeMets stopping rule. There was agreement to make the change. The second meeting was held in Barcelona. It was only necessary for me to spend two nights at the high-end hotel that the sponsor had chosen for this meeting. I wanted to spend a week in Barcelona and asked the sponsor to send me round trip tickets with departure and return on two consecutive Saturdays but only two nights at the meeting hotel. The sponsor made a reservation for me to stay 7 nights at this hotel at their expense and I told them I only wanted to stay two nights at the conference hotel at their expense. I would stay at another hotel at my expense for the remaining five days. The change was made. At the meeting I found that the protocol change, to correct the statistical error in interim analysis, that I had requested had not been made and, when I complained, the sponsor staff said they had no recollection of my request but they saw nothing wrong with the protocol. I continued to complain about the protocol but raised other questions about the safety analysis. Finally, the senior member of the sponsor team asked the group, "Don't you think we should invite our biostatistician member to the international renal cell carcinoma conference in Tokyo in October? We need another biostatistician there. Of course our company will pay all expenses." My fellow DMC members, all physicians who would be attending this conference on their own funds, surprisingly agreed. I did not respond to this offer. My fellow DMC members showed no interest in the protocol change I proposed even after I explained its statistical errors. After the meeting I wrote a letter indicating that if the protocol change was not made in 30 days, I would have to resign from this DMC. They made no reply and I sent a letter of resignation. A response from the sponsor accepted my resignation and indicated that I had a different idea of DMC responsibilities than they had. At an ASCO conference some years later the results of this trial were reported and, sure enough, a physician from the audience questioned the interim stopping rule and other shady aspects of the analysis. Have you ever heard of such a situation?

A No, you have had some novel experiences. I wonder what idea the sponsor had for the DMC responsibilities. The charter should have spelled this out. As you probably know, pharmaceutical companies have made headlines for offering trips to exotic places to practicing physicians as part of their marketing also claiming that the trip is to seek their valuable advice. This sponsor had not addressed your concerns on the protocol and your resignation was appropriate given the shady nature of this sponsor and lack of support

from your fellow DMC members. It would have been a big mistake for you to continue to serve, forget about the protocol, and accept the trip to Tokyo. There are some situations that can be fixed. This was not one of them.

Q8.2 I represent the sponsor of a clinical trial in obstetrics medicine. Some of our DMC members are spending a lot of time reading the tables we provide for their periodic meetings, having a lot of ad hoc phone conferences and now requiring further tables. This is increasing the cost of DMC operations beyond our budget and we do not see any significant safety issues on this trial. Are we hostage to this DMC because we cannot be trusted to review safety data ourselves?

A No, you are not hostage to the DMC. If the DMC is asking for more work, the chair should be able to justify this request without unmasking. You have a right to refuse the request or negotiate the request to a reasonable level. This does not take care of the problem of DMC members doing the routine review of tables but taking more hours than you had in your budget. One way to handle this problem would be to pay the DMC members a fixed quarterly fee regardless of number of meetings and the volume of the work done. Remind the DMC members that they are providing a service for drug development in their specialization. DMC service is not meant to be a lucrative consulting contract. However, if there is concern that circumstances are demanding more work than the quarterly fee will justify, a discussion between sponsor and DMC chair is appropriate with the latter having to justify these costs. After this conference the quarterly fee can be increased. This scheme can also be implemented by paying a fixed fee for each face-to-face meeting and teleconference regardless of the amount of preparation necessary.

Q8.3 I am a physician member of a DMC for a clinical trial for an experimental treatment for treatment-resistant tuberculosis. Our DMC just concluded a successful placebo-controlled phase II trial. FDA has granted accelerated approval for the U.S., but the sponsor is required to run a phase IV commitment controlled trial. The DMC and sponsor are in agreement that such a trial will not get the needed enrollment in the U.S. because patients there now have access to the newly approved drug through practitioners. The sponsor wants to do the trial in a third world country where they have no intention of marketing the product if approved. I feel this is immoral. Should I resign?

A There appears to be reason for you to resign on moral grounds but will that change the problem? It just removes you from participating. You and your fellow DMC members might get the sponsor to agree to market the drug in the third world country or do an active control trial where every patient will get a drug with some potential to help. The problem the sponsor is facing is inherent in the accelerated approval program. Your DMC could encourage the sponsor to talk to the FDA about this problem. There may be another way of meeting the phase IV commitment. The sponsor should definitely use your argument about the moral implications of doing a controlled trial of any kind in the third world when the sponsor has a recently approved drug for a serious disease. The sponsor's argument might also be aided by telling the FDA that the DMC cannot morally work on such a trial.

Q8.4 I am chair of a DMC working on an oncology trial for an Infant Pharma company. The DAC responsibilities are being handled by a CRO under contract with the sponsor. The DAC has now stopped work on this project because they are in a dispute regarding overdue payments from the sponsor. The DMC members are frustrated at being in the middle of this dispute and because of the uncertainty over when our next DMC meeting

can be held. One member has suggested that we all resign under the justification that "sponsor and DAC deserve one another." Should we do this? What options do we have?

A The first thing to consider when situations such as this arise is that the DMC exists to protect patient safety. The DMC should not resign unless suitable arrangements are made for the continuing safety surveillance. Money problems are a frequent concern with Infant Pharma but these sponsors know that they must work out their financial and contractor problems as soon as they can because the clinical trials are an important asset. The DAC is under contract with the sponsor but their contract is not like one for cleaning carpets or delivering bottled water. They too have a responsibility for patient safety, and they should carefully consider the implications of a work stoppage. Keep your DMC together. This too shall pass.

Appendix

TABLE A.1: Incidence of Selected Adverse Events for Approved Drugs in Placebo-Controlled Clinical Trials

Indication	Drug	Class/Type	Selected Adverse Events	Drug Incidence (%) [95% CI]	Placebo Incidence (%) [95% CI]	Odds Ratio [95% CI]
allergic rhinitis, seasonal	fexofenadine	antihistamine		$n = 679$	$n = 671$	
			viral infection	2.5 [1.5,4.0]	1.5 [0.7,2.7]	1.7 [0.7,4.2]
			nausea	1.6 [0.8,2.9]	1.5 [0.7,2.7]	1.1 [0.4,2.9]
			dysmenorrhea	1.5 [0.7,2.7]	0.3 [0.04,1.1]	5 [1.1,47.1]
			drowsiness	1.5 [0.7,2.7]	0.3 [0.04,1.1]	5 [1.1,47.1]
			dyspepsia	1.3 [0.6,2.5]	0.6 [0.2,2.5]	2.2 [0.6,10.0]
arthritis, rheumatoid	etanercept	TNF receptor blocker		$n = 349$	$n = 152$	
			injection site reaction	37.0 [31.8, 42.3]	10.0 [5.6, 15.8]	5.4 [3.0, 10.2]
			infection	35.0 [30.0,40.2]	32.0 [24.9, 40.3]	1.1 [0.7, 1.7]
			headache	17.0 [13.3, 21.3]	13.0 [8.8, 20.1]	1.3 [0.8, 2.5]
			rhinitis	12.0 [8.8, 15.9]	8.0 [4.1, 13.4]	1.6 [0.8, 3.4]
			nausea	9.0 [6.1,12.4]	10.0 [5.6, 15.8]	0.9 [0.5, 1.8]
cardiovascular	ramipril	ACE inhibitor		$n = 1004$	$n = 982$	
			hypotension	11.0 [9.1, 13.1]	5.0 [3.7, 6.5]	2.3 [1.7, 3.3]
			cough	8.0 [6.4, 19.8]	4.0 [2.8, 5.4]	2.1 [1.4, 3.1]
			dizziness	4.0 [2.9, 5.4]	3.0 {2.0, 4.2]	1.4 [0.8, 2.2]
			angina pectoris	3.0 [2.0, 4.2]	2.0 [1.2, 3.1]	1.5 [0.8, 2.6]
			nausea	2.0 [1.2, 3.1]	1.0 [0.5, 1.9]	2.0 [0.9, 4.2]
cardiovascular	metoprolol	beta blocker		$n = 700$	$n = 700$	
			hypotension	27.4 [24.2, 30.9]	23.2 [20.0, 26.5]	1.3 [1.0, 1.6]
			heart failure	27.5 [24.2, 30.9]	29.6 [26.2, 33.1]	0.9 [0.7, 1.1]
			bradycardia	15.9 [13.2, 18.8]	6.7 [5.0, 8.8]	2.6 [1.8, 3.8]
			heart blockage, 1^{st} degree	5.3 [3.7, 7.2]	1.9 [1.0, 3.2]	2.9 [1.5, 6.1]
			heart blockage, 2^{nd} or 3^{rd} degree	4.7 [3.3, 6.6]	4.7 [3.3, 6.6]	1.0 [0.6, 1.7]

depressive disorder, major	venlafaxine	antidepressant		$n = 357$	$n = 285$	
			insomnia	17.0 [13.3, 21.4]	11.0 [7.5, 15.1]	1.7 [1.0, 2.8]
			nervousness	10.0 [7.2, 13.7]	5.0 [2.7, 8.1]	2.2 [1.1, 4.4]
			anorexia	8.0 [5.5, 11.5]	2.0 [0.8, 4.5]	4.1 [1.6, 12.3]
			weight loss ($\geq$ 5%)	7.0 [4.8, 10.2]	2.0 [0.8, 4.5]	3.5 [1.4, 10.6]
	paroxetine	psychotropic		$n = 421$	$n = 421$	
			nausea	26.0 [21.8, 30.4]	9.0 [6.5, 12.2]	3.5 [2.3, 5.4]
			somnolence	23.0 [19.1, 27.4]	9.0 [6.5, 12.2]	3.0 [2.0, 4.6]
			asthenia	15.0 [11.7, 18.7]	6.0 [3.9, 8.6]	2.8 [1.7, 4.7]
			constipation	14.0 [10.8, 17.7]	9.0 [6.5, 12.2]	1.6 [1.0, 2.6]
			diarrhea	12.0 [9.2, 15.6]	8.0 [5.7, 11.1]	1.6 [1.0, 2.6]
diabetes mellitus, type II	rosiglitazone	thiazolidinedione		$n = 2526$	$n = 601$	
			upper respiratory infection	9.9 [8.8, 11.1]	8.7 [6.5, 11.2]	1.2 [0.8, 1.6]
			injury	7.6 [6.6, 8.7]	4.3 [2.8, 6.3]	1.8 [1.2, 2.9]
			headache	5.9 [5.0, 6.9]	5.0 [3.3, 7.1]	1.2 [0.8, 1.9]
			back pain	4.0 [3.3, 4.8]	3.8 [2.4, 5.7]	1.0 [0.7, 1.7]
			hyperglycemia	3.9 [3.2, 4.8]	5.7 [3.9, 7.8]	0.7 [0.5, 1.0]
duodenal ulcer, active	omeprazole	benzimidazole		$n = 465$	$n = 64$	
			headache	6.9 [4.8, 9.6]	6.3 [1.7, 15.2]	1.1 [0.4, 4.5]
			diarrhea	3.0 [1.7, 5.0]	3.1 [0.4, 10.8]	1.0 [0.2, 8.9]
			abdominal pain	2.4 [1.2, 4.2]	3.1 [0.4, 10.8]	0.8 [0.2, 7.1]
			nausea	2.2 [1.0, 3.9]	3.1 [0.4, 10.8]	0.7 [0.1, 6.5]
			upper respiratory infection	1.9 [0.9, 3.6]	1.6 [0.04, 8.4]	1.2 [0.2, 55.4]
epilepsy	oxcarbazepine	antiepileptic		$n = 126$	$n = 166$	
			vomiting	36.0 [27.4, 44.7]	5.0 [2.1, 9.3]	11.0 [4.8, 28.0]
			nausea	29.0 [21.6, 38.1]	10.0 [6.1, 15.9]	3.6 [1.9, 7.3]
			fatigue	15.0 [9.3, 22.5]	7.0 [3.8, 12.2]	2.3 [1.0, 5.4]
			abdominal pain	11.0 [6.2, 17.9]	5.0 [2.1, 9.3]	2.5 [0.9, 7.0]
			diarrhea	7.0 [3.3, 13.1]	6.0 [2.9, 10.8]	1.2 [0.4, 3.4]

(*Continued*)

TABLE A.1: Incidence of Selected Adverse Events for Approved Drugs in Placebo-Controlled Clinical Trials (Continued)

Indication	Drug	Class/Type	Selected Adverse Events	Drug Incidence (%) [95% CI]	Placebo Incidence (%) [95% CI]	Odds Ratio [95% CI]
erectile dysfunction	sildenafil	sexual dysfunction		$n = 734$	$n = 725$	
			headache	16.0 [13.4, 18.8]	4.0 [2.7, 5.7]	4.6 [3.0, 7.2]
			flushing	10.0 [7.8, 12.3]	1.0 [0.4, 2.0]	11.3 [5.2, 29.3]
			dyspepsia	7.0 [5.2, 9.0]	2.0 [1.2, 3.4]	3.5 [1.9, 6.8]
			nasal congestion	4.0 [2.7, 5.6]	2.0 [1.2, 3.4]	1.9 [1.0, 3.9]
			urinary tract infection	3.0 [1.9, 4.5]	2.0 [1.2, 3.4]	1.5 [0.7, 3.1]
fibromyalgia	pregabalin	neurotransmitter		$n = 600$	$n = 505$	
			dizziness	45.0 [41.0, 49.1]	9.0 [6.6, 11.7]	8.3 [5.9, 12.1]
			somnolence	22.0 [18.8, 25.5]	4.0 [2.4, 6.1]	6.8 [4.2, 11.8]
			weight increase	14.0 [11.3, 17.0]	2.0 [1.6, 3.6]	8.1 [4.1, 17.6]
			vision, blurred	12.0 [9.5, 14.9]	1.0 [0.3, 23.0]	13.6 [5.5, 43.6]
			edema, peripheral	9.0 [6.8, 11.6]	2.0 [1.0, 3.6]	4.9 [2.4, 10.9]
			dry mouth	9.0 [6.8, 11.6]	2.0 [1.0, 3.6]	4.9 [2.4, 10.9]
hypercholesterolemia[a]	atorvastatin	lipid lowering		$n = 863$	$n = 270$	
			infection	10.3 [8.4, 12.5]	10.0 [6.7, 14.2]	1.0 [0.6, 1.7]
			headache	5.4 [4.0, 7.2]	7.0 [4.3, 10.8]	0.8 [0.4, 1.4]
			rash	3.9 [2.7, 5.5]	0.7 [0.1, 2.7]	5.5 [1.4, 47.5]
			abdominal pain	2.8 [1.8, 4.1]	0.7 [0.1, 2.7]	3.8 [0.9, 33.7]
			back pain	2.8 [1.8, 4.1]	3.0 [1.3, 5.8]	0.9 [0.4, 2.4]
osteoporosis	risedronate	bisphosphonate		$n = 1914$	$n = 1916$	
			arthralgia	21.1 [19.3, 23.0]	23.7 [21.8, 25.7]	0.9 [0.7, 1.0]
			abdominal pain	11.6 [10.2, 13.1]	9.4 [8.1, 10.8]	1.3 [1.0, 1.6]
			urinary tract infection	10.9 [9.5, 12.4]	9.7 [8.4, 11.1]	1.1 [0.9, 1.4]
			hypertension	10.0 [8.7, 11.4]	9.0 [7.7, 10.4]	1.1 [0.9, 1.4]
			joint disorder	6.8 [5.7, 8.0]	5.4 [4.4, 6.5]	1.3 [1.0, 1.7]

	teriparatide	recombinant human parathyroid hormone		$n = 691$	$n = 691$	
			arthralgia	10.1 [8.0, 12.6]	8.4 [6.4, 10.7]	1.2 [0.8, 1.8]
			rhinitis	9.6 [7.5, 12.0]	8.8 [6.8, 11.2]	1.1 [0.7, 1.6]
			nausea	8.5 [6.4, 10.7]	6.7 [4.9, 8.8]	1.3 [0.8, 2.0]
			dizziness	8.0 [6.1, 10.2]	5 [3.8, 7.3]	1.5 [1.0, 2.4]
			hypertension	7.1 [5.3, 10.3]	6.8 [5.0, 8.9]	1.0 [0.7, 1.6]
platelet reduction	clopidogrel	adenosine inhibitor		$n = 9599$	$n = 9586$[a]	
			headache	7.6 [7.1, 8.1]	7.2 [6.9, 7.7]	1.1 [1.0, 1.2]
			arthralgia	6.3 [5.8, 6.8]	6.2 [5.7, 6.7]	1.0 [0.9, 1.1]
			diarrhea	4.5 [4.1, 4.9]	3.4 [3.0, 3.8]	1.3 [1.2, 1.6]
			hypertension	4.3 [3.9, 4.7]	5.1 [4.7, 5.6]	0.8 [0.7, 1.0]
			nausea	3.4 [3.0, 3.8]	3.8 [3.4, 4.2]	0.9 [0.8, 1.0]
renal failure, chronic	epoetin	glycoprotein		$n = 200$	$n = 135$	
			hypertension	24.0 [18.3, 30.5]	19.0 [13.0, 26.9]	1.3 [0.8, 2.4]
			headache	16.0 [11.2, 21.8]	12.0 [6.9, 18.5]	1.4 [0.7, 2.9]
			arthralgia	11.0 [7.0, 16.2]	6.0 [2.6, 11.3]	2.0 [0.8, 5.3]
			nausea	11.0 [7.0, 16.2]	9.0 [4.7, 15.0]	1.3 [0.6, 2.9]
			edema	9.0 [5.4, 13.9]	10.0 [5.8, 16.8]	0.9 [0.4, 1.9]
sleep disorder	eszopiclone	non-benzodiazepine hypnotic agent		$n = 105$	$n = 99$	
			unpleasant taste	34.0 [25.3, 44.2]	3.0 [0.6, 8.6]	16.7 [4.9, 87.1]
			headache	17.0 [10.5, 25.7]	13.0 [7.2, 21.4]	1.4 [0.6, 3.2]
			respiratory infection	10.0 [5.3, 18.0]	3.0 [0.6, 8.6]	3.7 [0.9, 21.4]
			somnolence	8.0 [3.3, 14.5]	3.0 [0.6, 8.6]	2.6 [0.6, 15.8]
			dizziness	7.0 [2.7, 13.3]	4.0 [1.1, 10.0]	1.7 [0.4, 8.1]

Note: This material is presented to illustrate material in the text. It is not a substitute for official professional and consumer information on these drugs.

[a] Control group received aspirin, not placebo.

Source: (raw data) U.S. Food and Drug Administration (2008), Drugs@FDA, http://www.accessdata.fda.gov/scripts/cder/drugsatfda/index.cfm.

Glossary

(Terms are defined within the context of this book rather than more generally.)

active control—the control treatment is an approved drug rather than a placebo (Chapter 6).

adaptive design—a clinical trial design where the trial becomes redesigned due to data collected during the trial, such as changing sample size, dropping a dose group (Chapter 8).

add-on design—a clinical trial design that compares treatment A with treatment A + B where A is an approved drug and B is the experimental drug (Chapter 6).

ad hoc consultant—a person appointed to assist a DMC on specific matters involving expertise not present on the DMC such as an allergist being consulted on hypersensitivity on an ophthalmology trial (Chapter 2).

ad hoc meeting—a DMC meeting that had not been previously scheduled but is called to discuss a specific, recently occurring safety issue (Chapter 3).

adherers—those patients enrolled in a trial who are complying with the protocol (Chapter 5).

adverse event (AE)—An *adverse event* is any unfavorable and unintended sign (including an abnormal laboratory finding, for example), symptom, or disease temporally associated with the use of a drug, whether or not considered related to the drug. The term *treatment emergent* is often added as a modifier in order to remove manifestations of preexisting conditions from consideration (Chapter 4).

adverse event grade—see **adverse event severity.**

adverse event severity—a measure of the intensity or extent of the event, sometimes called grade (Chapter 4).

alternative hypothesis—in statistical hypothesis testing the hypothesis that indicates that there is a difference in adverse event incidence between treatment groups (Chapter 5).

APPROVe—acronym for the Adenomatous Polyp Prevention on Vioxx clinical trial (Chapter 5).

assay sensitivity—a statistical calculation of the magnitude of treatment difference in adverse event incidence that can be detected with the sample sizes in each treatment group (Chapter 5).

attained significance level—see ***p*-value.**

bad news travels first—a phenomenon in clinical trial communications where data reporting deaths or SAEs arrive at the sponsor before data on routine visits, thus putting these events out of proportion to follow-up time on the trial (Chapter 7).

Bayesian methods—statistical methods that combine prior information on adverse event incidence with data collected on the clinical trial (Chapter 5).

bias—an observed treatment difference due to effects other than treatments themselves and/or nonobjective actions in operations or evaluations (Chapter 6).

Big Pharma—a pharmaceutical firm with more than $8 billion in revenues in 2006 and many products on the market (Chapter 1).

binomial distribution—a statistical distribution used to determine the probability of observing a specified number of adverse events given an assumed adverse event incidence (Chapter 5).

biomarkers—a clinical measurement that is predictive of a future outcome such as a skin rash predicting efficacy in an oncology trial (Chapter 8).

black box warning—a communication between a regulatory agency and practicing physicians of a very important serious adverse event associated with a drug; the warning is enclosed in a black box on the package insert (Chapter 1).

blinded—treatment identity is hidden; see also **masked** (Chapter 1).

case report form (CRF)—a form used to collect all of the clinical data during the trial. It should be reviewed by the DMC prior to the outset of the trial (Chapter 4).

causal inference—a statistical method that allows conclusions of causation in clinical trials for variables other than treatment group (Chapter 8).

censored observation—in time-to-event analysis the time contribution for a patient who did not experience an adverse event (Chapter 5).

chi-square test—a statistical method for assessing the difference in adverse event incidence between treatment groups (Chapter 5).

CIOMS form—a form developed by the Council for International Organizations of Medical Sciences that is used for narrative reports of adverse events. It ensures that the history, concomitant medications, preexisting conditions, comorbidity, and so on are all reported. Copies of CIOMS forms, or a sponsor's own version, are distributed to DMC members for all SAEs (Chapter 4).

clinical research associate (CRA)—a member of the safety monitoring team who provides site monitoring services (Chapter 1).

Clopper–Pearson confidence interval—see **exact binomial confidence interval.**

closed session—the phase of a DMC meeting where the DMC members and the DAC statistician meet (Chapter 3).

competing risks—a statistical artifact that occurs when the occurrence of one event causes a reduced likelihood of another event; in the context of safety analysis a treatment group where patients experience early deaths may have less cardiovascular toxicity than a treatment group where patients don't die early, but this might just be due to the fact that patients who die early are not treated long enough to develop cardiovascular adverse events (Chapter 6).

conditional power—given the data at hand at an interim point in a clinical trial, the probability of detecting a treatment difference in adverse event incidence at the end of the trial under the assumption of the trajectory of future data (Chapter 5).

confidence interval—a statistical method for estimating a plausible range for adverse event incidence (Chapter 5).

confirmatory trial—typically a phase III clinical trial. It is the trial(s) on which regulatory approval will be based. Also called a pivotal trial. Most DMCs are working on confirmatory trials (Chapter 1).

conflict of interest—a situation where a DMC member might not be considered sufficiently independent such as if he/she owns stock in the sponsor or a company offering a competing product; this possibility must be reviewed before a member may be appointed to a DMC (Chapter 2).

contract research organization (CRO)—an organization under contract with the sponsor to take over tasks often performed by the sponsor such as site monitoring, report writing, biostatistics (Chapter 2).

covariate—a variable that might affect AE incidence such as treatment or geographic region (Chapter 5).

CTCAE (Common Terminology Criteria for Adverse Events)—a classification system like MedDRA developed by the U.S. National Cancer Institute for classifying adverse events in oncology trials sponsored by NCI; oncology investigators must be trained to use MedDRA in pharmaceutical industry–sponsored trials (Chapter 4).

data analysis center (DAC)—the organization with the responsibility for preparing tables, listings, graphs, analyses, and so on for the DMC. The DAC will be unmasked so it would have to be a separate independent unit if within the sponsor. The DAC is often found at a contract research organization working under contract with the sponsor. The DAC contributes an independent statistician as a nonvoting member of the DMC (Chapter 2).

data monitoring committee (DMC)—a committee created by a sponsor to provide independent review of accumulating safety and efficacy data. This committee is responsible for the stewardship of the trial which translates to protecting trial integrity and patient safety (Chapter 1).

data review meeting—a regular occurring DMC meeting where the DMC will review accumulated safety data (Chapter 3).

"Dear Investigator" letter—a letter written by a sponsor to investigators alerting them to unexpected adverse events observed on the trial or asking them to change operations for supportive care, infection control, and so on (Chapter 7).

decision matrix—a table created by a DMC in closed session prior to unmasking indicating the course of action to be followed for various potential distributions of adverse events between treatment groups (Chapter 7).

deductive inference—in medicine, given the disease what symptoms can be expected (Chapter 5).

DMC charter—a document approved by the sponsor and DMC at the outset of a clinical trial that indicates the scope and rules and regulations of DMC operations (Chapter 2).

double false discovery rate (DFDR)—a method of controlling for multiplicity by applying the false discovery rate (FDR) twice (Chapter 5).

drug label—see **package insert.**

effect size—a measure of the difference in an efficacy parameter between experimental and control treatment (Chapter 8).

ethics committee—see **institutional review board.**

evidence—information gained from a clinical trial on adverse event incidence (Chapter 5).

exact binomial confidence interval—a confidence interval computed on the basis of the binomial distribution rather than the normal distribution (Chapter 5).

executive session—if needed, the phase of a closed session where the DMC members meet without the presence of the DAC statistician (Chapter 3).

expedited SAE—an SAE that is unexpected; regulatory agencies require prompt reporting of this class of SAEs, and they should simultaneously be reported to the DMC (Chapter 4).

exploratory trials—typically phase I and II clinical trials (Chapter 1).

false discovery rate (FDR)—a statistical method for controlling multiplicity by considering the proportion of all hypotheses declared statistically significant where in fact no treatment effect existed (Chapter 5).

false negative—in statistical hypothesis testing, concluding that a treatment difference does not exist when, in fact, a difference exists (Chapter 5).

false positive—in statistical hypothesis testing, concluding a treatment difference when, in fact, the difference does not exist (Chapter 5).

familywise error rate (FWER)—see **Type I error.**

frequentist methods—traditional statistical methods based on repeated sampling (Chapter 5).

firewall—an internal sponsor process by which staff members working on the clinical trial remain masked (Chapter 2).

firewall committee—see **pharmacovigilence committee.**

Fisher's exact test—a statistical method similar to a chi-square test but using a discrete probability distribution instead of the normal distribution to determine the *p*-value (Chapter 5).

futility—a result of an interim efficacy analysis that indicates that it is not likely that a statistically significant treatment difference will be found if the trial is to continue to completion (Chapter 7).

grade—see **adverse event severity.**

granularity—a phenomenon that arises in adverse event classification indicating the extent of specificity in preferred terms (Chapter 6).

hazard ratio—in time-to-event analysis, the ratio of risk of adverse event per unit of time in the experimental group to the control group (Chapter 5).

incidence, percent—proportional incidence $\times$ 100% (Chapter 5).

incidence, proportional—the ratio of the number of patients experiencing an adverse event to the total number of patients randomized to that treatment group (Chapter 5).

incomplete observation—see **censored observation.**

indemnification—a provision in a DMC member's contract whereby the sponsor pledges to protect DMC members against liability issues that may arise in the trial that are not their fault (Chapter 2).

inductive incidence—in medicine, given the symptoms what disease can we expect (Chapter 5)?

independent statistician—a statistician employed by the DAC who will serve as a nonvoting member of the DMC and unmask the DMC members if requested (Chapter 2).

Infant Pharma—a pharmaceutical firm with no revenues from product sales and no products on the market (Chapter 1).

informed consent—a document prepared by the institution performing a clinical trial which informs potential patient volunteers of, among other things, possible adverse events they might experience in the trial (Chapter 3).

institutional review board (IRB)—in the United States a committee that exists at every institution performing research on human subjects. The committee reviews all aspects of the research in order to protect the safety of human subjects. Similar committees in other countries are called ethics committees (Chapter 2).

integrated summary of safety (ISS)—a section of a new drug application that summarizes safety data for the experimental treatment over all human trials. It should

be established at the DMC orientation meeting if the DMC will be expected to review the ISS prior to submission to the regulatory agency (Chapter 3).

internal safety review committee—a group of sponsor employees not working on the clinical trial of interest performing a function in parallel with a DMC (Chapter 8).

International Conference on Harmonisation (ICH)—a joint effort among the United States, Europe, and Japan to create international agreement on regulations for the approval of new drugs (Chapter 1).

investigator brochure—a document prepared by the sponsor before the trial begins that summarizes all known information about the experimental drug used in the trial; it should be reviewed by the DMC at the outset of a trial (Chapter 3).

investigator-sponsored trial—an experimental clinical trial where the investigator takes on all responsibility for reporting progress and adverse events to a regulatory agency (Chapter 1).

Kaplan–Meier graph—a statistical method of graphically depicting the cumulative frequency of adverse events over time (Chapter 5).

lambda (λ)—the parameter (mean) of the Poisson distribution (Chapter 5).

landmark estimate—in time-to-event analysis the estimate of adverse event incidence at a particular point in time such as 12 months after treatment start (Chapter 5).

learning trials—see **exploratory trials.**

likelihood function—an expression giving the degree of belief in various levels of adverse event incidence conditional on the data observed in the trial (Chapter 5).

likelihood graph—a graph showing the relative likelihood for each value of adverse event incidence (Chapter 5).

likelihood methods—statistical methods based on the data already collected rather than repeated sampling (Chapter 5).

log rank test—a statistical method for determining the *p*-value for the difference between treatment groups in a time-to-event/Kaplan–Meier analysis (Chapter 5).

logistic regression analysis—a multifactor method to assess the influence of covariate factors on AE incidence and compute adjusted rates (Chapter 5).

masked—treatment is hidden; synonym for **blinded** but used here to avoid confusion in ophthalmology clinical trials (Chapter 1).

masked, partially—a policy whereby DMC members know the treatments as A, B, C, and so on, but do not know the identity of the treatment codes (Chapter 5).

MedDRA—medical dictionary for regulatory affairs; a system for coding adverse events that is generally used in pharmaceutical industry clinical trials (Chapter 4).

medical governance committee—see **pharmacovigilence committee.**

meta-analysis—a statistical method for bringing together data from previous clinical trials in an effort to create one combined estimate of effect size or adverse event incidence (Chapter 7).

meta-analysis, fixed effects—the treatment effect being estimated is assumed to be fixed across trials (Chapter 7).

meta-analysis, prospective—performed on raw data from various trials (Chapter 7).

meta-analysis, random effects—the treatment effect being estimated is assumed to vary randomly across trials (Chapter 7).

meta-analysis, retrospective—performed on data extracted from the literature (Chapter 7).

Middle Pharma—a pharmaceutical firm with less than $8 billion in revenues in 2006 and some products on the market (Chapter 1).

monitoring, safety—continual review of accumulating safety data during a clinical trial (Chapter 1).

monitoring, site—a quality control procedure applied periodically during the trial by sponsor or contract clinical research associates (Chapter 1).

monitoring, statistical—making calculations on accumulating efficacy data to justify early termination of a clinical trial (Chapter 1).

multiplicity—the statistical phenomenon where many tests of a hypothesis increase the probability of false positive results (Chapter 5).

noninferiority trial—a clinical trial that has its objective to show that an experimental treatment is no worse than active control and possibly superior (Chapter 8).

null hypothesis—in statistical hypothesis testing the hypothesis that there is no difference in adverse event incidence between treatment groups (Chapter 5).

odds ratio—a measure of relative risk of adverse event in experimental treatment group to control group (Chapter 5).

one-sided test—a statistical hypothesis test that considers whether adverse event incidence in the experimental treatment group is statistically significantly greater than in control group (Chapter 5).

open label—treatment identity is known; opposite of **masked** (Chapter 1).

open label extension studies—a follow-up program where patients exit a clinical trial but continue on the experimental treatment for an uncontrolled phase in order to gather long-term safety data; often used in neurology and oncology (Chapter 1).

open session—the first phase of a DMC meeting where the sponsor staff, DAC members, and DMC members meet to discuss trial progress and report on issues of interest to all (Chapter 3).

orientation (organizational) meeting—the first meeting between the DMC, the DAC and the sponsor staff (Chapter 3).

***p*-value**—the probability that a treatment difference could have occurred due to chance if, in fact, there was no population difference in treatments (Chapter 5).

package insert—sometimes called the drug label, the document prepared by the sponsor that provides all prescribing information to the physician including adverse events. It should be established at the DMC orientation meeting whether or not the DMC will be responsible for reviewing the proposed package insert prior to submission to the regulatory agency (Chapter 3).

pharmacovigilence staff—a group of sponsor employees separated from the clinical trial staff by a "firewall" charged with reviewing adverse events, possibly unmasked, during the trial; they may call matters of concern to the attention of the DMC (Chapter 4).

pharmacovigilence committee—a committee of sponsor staff not involved in the clinical trial with whom the DMC can consult with if they have a serious safety concern. This avoids unmasking the sponsor staff working on the trial. It is sometimes called a firewall committee or medical governance committee (Chapter 3).

phase I trial—early safety trial; DMCs usually not involved (Chapter 1).

phase II trial—safety trial to refine dose and early efficacy trial; DMCs sometimes involved (Chapter 1).

pivotal trial—the trial(s) on which regulatory approval will be based, also called a confirmatory trial, typically phase III; most DMCs are working on pivotal trials (Chapter 1).

Poisson distribution—a probability distribution for rare adverse events based on a mean per time unit (Chapter 5).

Poisson rate ratio—the ratio of rate per 100 patient years in experimental treatment group to control group (Chapter 5).

posterior distribution—in Bayesian statistical analysis, the probability distribution of adverse event incidence that results when the prior distribution is combined with the distribution of data collected in the clinical trial (Chapter 5).

postmarket surveillance—the process of monitoring incidence of adverse events after a drug has been approved, usually performed by forms' being submitted to the sponsor by practicing physicians (Chapter 1).

power—in statistical hypothesis testing the probability that the null hypothesis of no treatment difference will be rejected when the treatments, in fact, differ given the sample size in each group (Chapter 5).

predictive power—the Bayesian analog to conditional power (Chapter 5).

preferred term (PT)—in MedDRA a term within a system order class (SOC) to describe a specific adverse event (e.g., within cardiovascular, myocardial infarct) (Chapter 4).

progression-free survival—in oncology trials a patient is defined as a treatment failure if he/she experiences progression of disease or dies from any cause, whichever occurs first (Chapter 6).

prior distribution—in Bayesian statistical analysis a probability distribution describing prior beliefs in adverse event incidence distribution (Chapter 5).

proof of concept—a term usually used in patent law used here to describe a preclinical study or an early phase clinical trial (Chapter 1).

publication bias—a phenomenon that occurs when a clinical trial with a positive result is more likely to be published than one with negative results (Chapter 7).

rate per 100 patient years—computed as the number of patients experiencing and adverse event $\times$ 100 divided by the total patient years contributed by that treatment group (Chapter 5).

reconsenting—a process whereby, due to the discovery of unexpected adverse events, patients are asked to sign a revised informed consent form (Chapter 7).

run-in screening phase—a prerandomization phase of a clinical trial where prospective patients must qualify for the trial by, for example, recording a sufficient number of seizures or having a diastolic blood pressure within a certain range (Chapter 5).

safety monitoring—see **monitoring, safety.**

safety monitoring plan (SMP)—a document prepared by the sponsor at the outset of a clinical trial indicating responsibilities and procedures for the DMC, DAC, sponsor pharmacovigilence staff, investigators, clinical research associates, and so on; it should be reviewed by the DMC at the outset of a trial (Chapter 3).

safety monitoring team—a term used to encompass all persons and organizations involved with the monitoring of safety during a clinical trial; it would include the sponsor clinical research and pharmacovigilence staffs, CRO staff, DMC members, DAC, and so on (Chapter 3).

screening phase—see **run-in screening phase.**

seamless transition, phase II to phase III—an adaptive design where a phase II trial becomes a phase III trial after certain conditions are satisfied and the phase II data are used as part of the phase III trial (Chapter 8).

serious adverse event (SAE)—any untoward medical occurrence that, at any dose, results in death, is life-threatening, requires inpatient hospitalization or prolongation of existing hospitalization, results in persistent or significant disability/incapacity, or is a congenital anomaly/birth defect (Chapter 4).

severity—see **adverse event severity.**

significance level—a pretrial setting of the p-value that will determine statistical significance, often p-value less than 0.05 (Chapter 5).

site monitoring—see **monitoring, site.**

sponsor—the organization that has the ultimate responsibility for reporting the results to the regulatory authorities. For our purposes it will most often be a pharmaceutical or biotechnology company but it could be a university, government agency, or in the case of orphan drugs, a patient–parent support group (Chapter 1).

sponsor representative—an employee of the sponsor who will serve as the liaison with the DMC (Chapter 2).

statistical monitoring—see **monitoring, statistical.**

standard operating procedures (SOPs)—written procedures used by a sponsor covering all clinical operations; this should include procedures for the formation and operation of a DMC (Chapter 2).

statistical analysis plan (SAP)—a document prepared before the start of a clinical trial indicating all definitions, assumptions, and methods for the analysis of data at the conclusion of the trial (Chapter 3).

stewardship—the act of providing careful and responsible management of something entrusted in one's care; used here to describe the overall responsibility of a DMC (Chapter 1).

superiority trial—a clinical trial whose objective is to show that the experimental treatment is superior in efficacy to control (Chapter 8).

support limits—the likelihood analog of confidence interval, a range of plausible values of adverse event incidence or relative risk (Chapter 5).

system organ class (SOC)—In MedDRA, a hierarchical adverse event classification system, the SOC is the highest order—gastrointestinal, nervous system, cardiovascular system, and so on (Chapter 4).

test of concept—see **proof of concept.**

theta (θ)—the Poisson rate ratio (Chapter 5).

time-to-event analysis—a statistical method that takes into consideration the time to the onset of adverse events. See also **Kaplan–Meier graph** (Chapter 5).

treatment-emergent—a modifier for adverse events used to remove manifestations of preexisting conditions from consideration (Chapter 4).

two-sided test—a statistical hypothesis test that considers whether adverse event incidence in the experimental treatment group is either statistically significantly greater than control or statistically significantly less than control (Chapter 5).

Type I error—in statistical hypothesis testing, the probability that the null hypothesis of no treatment difference is rejected when the null hypothesis is true (Chapter 5).

Type II error—the power or the probability that the null hypothesis is not rejected when it is in fact false (Chapter 5).

unmasking, de facto—persons working on a clinical trial are unmasked to treatment group by observing differences in adverse event frequency, differential visit schedules or different procedures at visits between treatment groups (Chapter 6).

unmasking, deliberate—a DMC asks the independent statistician to unmask them (Chapter 7).

white space—a term used by adaptive design planners to indicate a period of time when analysis and planning takes place prior to an adaptation of the trial (Chapter 8).

List of Abbreviations

AE—adverse event
ALA—Applied Logic Associates
APPROVe—Adenomatous Polyp Prevention on Vioxx
CDC—Center for Disease Control and Prevention
CIOMS—Council of International Organizations of Medical Sciences
CRA—clinical research associate
CRF—case report form
CRO—contract research organization
CTCAE—Common Terminology Criteria for Adverse Events
DAC—Data Analysis Center
DFDR—Double False Discovery Rate
DMC—Data Monitoring Committee
FDA—Food and Drug Administration
FDR—false discovery rate
FWER—family-wise error rate
HDL—high-density lipoprotein
HIV—human immunodeficiency virus
ICH—International Conference on Harmonisation
IFPMA—International Federation of Pharmaceutical Manufacturers and Associations
IQWiG—German Institute for Quality and Efficiency in Health Care
IRB—Institutional Review Board
ISRC—Internal Safety Review Committee
ISS—integrated summary of safety
ITE—insufficient therapeutic effect
LDL—low density lipoprotein
LRC—lipids research clinics
MD—doctor of medicine
MedDRA—Medical Dictionary for Regulatory Affairs
MSSO—Maintenance Support and Service Organization
NICE—National Institute of Health and Clinical Excellence
NIH—National Institutes of Health
NSAID—nonsteroidal anti-inflammatory drug
OR—odds ratio
PFS—progression-free survival
PharmD—doctor of pharmacy

PhD—doctor of philosophy
QTc—in cardiology, heart rate-corrected QT interval
R&D—research and development
RR—relative risk
SAE—serious adverse event
SAP—statistical analysis plan
SMP—safety monitoring plan
SOC—system organ class
SOP—standard operating procedure
UK—United Kingdom
U.S.—United States
VA—Veterans Administration
WHO—World Health Organization

References

Agresti, A. (1992) A survey of exact inference for contingency tables, *Statistical Sciences*, **7**, 131–177.

Aronson, J.K. and Ferner, R.E. (2005) Clarification of terminology in drug safety, *Drug Safety*, **28**, 851–870.

Ashcroft, D., Li Wan Po, A., Williams, H. et al (1999) Clinical measures of disease severity and outcome in psoriasis: a critical appraisal of their quality, *British Journal of Dermatology*, **141**, 185–191.

Baron, J.A., Sandler, R.S., Bresalier, R.S. et al. (2006) A randomized trial of rofecoxib for the chemoprevention of colorectal adenomas, *Gastroenterology*, **131**, 1674–1682.

Benjamini, Y. and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing, *Journal of the Royal Statistical Society, Series B*, **57**, 289–300.

Berlin, J.A. (2008) Use of meta-analysis in drug safety assessments. Paper presented at Spring Statistical Meetings, Eastern North American Region, International Biometrics Society, Arlington, VA.

Berry, S.M. and Berry, D.A. (2004) Accounting for multiplicities in assessing drug safety: a three-level hierarchical mixture model, *Biometrics*, **60**, 418–426.

Berry, D.A. (2005) Introduction to Bayesian methods III: use and interpretation of Bayesian tools in design and analysis, *Clinical Trials*, **2**, 295–300.

Blanchard, K., Carlos, J.P., and Randolph, A. (2001) *The Three Keys to Empowerment*, San Francisco: Barrett-Koehler.

Blume, J.D. (2002) Tutorial in biostatistics: likelihood methods for measuring statistical evidence, *Statistics in Medicine*, **21**, 2563–2599.

Bombardier, C., Gladman, D.D., and Urowitz, M.B. (1992) Derivation of the SLEDAI: A disease activity index for lupus patients, *Arthritis and Rheumatism*, **35**, 630–640.

Bonhoeffer, J., Kohl, K., Chen, R. et al. (2002) The Brighton Collaboration addressing the need for standardized case definitions of adverse events following immunization (AEFI), *Vaccine*, **21**, 298–302.

Bousquet, C., Lagier, G., and Lillo-Le Louet, A. (2005) Appraisal of the MedDRA conceptual structure for describing and grouping adverse drug reactions, *Drug Safety*, **28**, 19–34.

Bresalier, R.S., Sandler, R.S., Quan, H. et al. (2005) Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial, *New England Journal of Medicine*, **352**, 1092–1102.

Brown, E.G., Wood, L., and Wood, S. (1999) The medical dictionary for regulatory activities (MedDRA), *Drug Safety*, **20**, 109–117.

Bush, J.K., Dai, W.S., Dieck, G.S. et al. (2005) The art and science of risk management: a US research-based industry perspective, *Drug Safety*, **28**, 1–18.

Carpenter, D., Zucker, E.J., and Avorn, J. (2008) Drug review deadlines and safety problems, *New England Journal of Medicine*, **358**, 1354–1361.

Charney, D.S., Nemeroff, C.B., Lewis, L. et al. (2002) National Depressive and Manic-Depressive Association consensus statement on the use of placebo in clinical trials of mood disorders, *Archives of General Psychiatry*, **59**, 262–270.

Chuang-Stein, C., Anderson, K., Gallo, P. et al. (2006) Sample size reestimation: a review and recommendations, *Drug Information Journal*, **40**, 475–484.

Cleves, M., Gould, W., Guttierrez, R. et al. (2008) *An Introduction to Survival Analysis Using Stata*, Second Ed., Boca Raton, FL: CRC Press.

Clopper, C.J. and Pearson, E.S. (1934) The use of confidence or fiducial limits illustrated in the case of the binomial, *Biometrika*, **26**, 404–413.

Code of Federal Regulations (2005) *Protection of Human Subjects -45 CFR Part 46*, http://ohsr.od.nih.gov/guidelines/45cfr46.html.

Cohen, S.B., Emery, P., Greenwald, M.W. et al. (2006) Rituximab for rheumatoid arthritis refractory to anti-tumor necrosis factor therapy: results of a multicenter, randomized, double blind placebo-controlled Phase III trial evaluating primary efficacy and safety at 24 weeks, *Arthritis and Rheumatism*, **54**, 2793–2806.

Cornfield, J. (1973) Remarks made at Spring Biostatistics Meeting, Ithaca, NY.

Cox, D.R. (1972) Regression models and life tables (with discussion), *Journal of the Royal Statistical Society B*, **34**, 187–220.

Cranney, A. and Adachi, J.D. (2005) Benefit-risk assessment of raloxifene in postmenopausal osteoporosis. *Drug Safety*, **28**, 721–730.

Cui, L, Hung, M.H.J., and Wang, S.J. (1999) Modification of sample size in group sequential clinical trials, *Biometrics*, **55**, 853–857.

Cytel (2007) *StatXact 8 User Manual*, Cambridge, MA: Cytel, Inc.

Day, R.O. and Williams, K.M. (2007) Open-label extension studies: do they provide meaningful information on the safety of new drugs? *Drug Safety*, **30**, 93–105.

D'Agostino, R.B., Massaro, J.M., and Sullivan, L.M. (2003) Non-inferiority trials: design concepts and issues—the encounters of academic consultants in statistics, *Statistics in Medicine*, **22**, 169–186.

DeMets D.L., Fleming T.R., Rockhold, F. et al. (2004) Liability issues for data monitoring committee members, *Clinical Trials*, **1**, 525–531.

DeMets D.L., Furberg C.D., and Friedman, L.M. (2006) *Data Monitoring in Clinical Trials: A Case Studies Approach*, New York: Springer.

DerSimonian, R. and Laird, N. (1986) Meta-analysis in clinical trials, *Controlled Clinical Trials*, **7**, 177–178.

Devchand, P.R. (2008) Glitazones and the cardiovascular system, *Current Opinion in Endocrinology, Diabetes and Obesity*, **15**, 188–192.

Dickersin, K., Olson, C.M., Rennie, D. et al. (2002) Association between time interval to publication and statistical significance: reports of controlled trials published in JAMA, *JAMA* **287**, 2829–2831.

Dickersin, K. and Rennie, D. (2003) Registering clinical trials, *JAMA* **290**, 516–523.

Dixon, D.O., Freedman, R.S., Herson, J. et al. (2006) Guidelines for data and safety monitoring for clinical trials not requiring traditional data monitoring committees, *Clinical Trials*, **3**, 314–319.

Djunaedi, E., Sommer, A., Pandji, A., and the ACEH Study Group (1988) Impact of vitamin A supplementation on xerophthalmia. A randomized controlled community trial. *Archives of Ophthalmology*, **106**, 218–222.

Dragalin, V. (2006) Adaptive designs: terminology and classification, *Drug Information Journal*, **40**, 425–435.

Edwards, I.R. and Biriell, C. (1994) Harmonisation in pharmacovigilance, *Drug Safety*, **10**, 93–102.

Ellenberg, S.S., Fleming, T.R., and DeMets, D.L. (2002) *Data Monitoring Committees in Clinical Trials: A Practical Perspective*, New York: John Wiley & Sons.

Ellenberg, S.S. and George, S.L. (2004) Should statisticians reporting to data monitoring committees be independent of the trial sponsor and leadership? *Statistics in Medicine*, **23**, 1503–1505.

Ellis, A.J., Hughes, R.D., Wendon, J.A. et al. (1996) Pilot-controlled trial of extracorporeal liver assist device in acute liver failure, *Hepatology*, **24**, 1446–1451.

Fairweather, W.R. (1996) Integrated safety analysis: statistical issues in the assessment of safety in clinical trials, *Drug Information Journal*, **30**, 875–879.

Faught, E., Sachdeo, R.C., Remler, M.P. et al. (1993) Felbamate monotherapy for partial onset seizures: an active control trial, *Neurology*, **43**, 688–692.

Fay, M.P., Huang, C.-Y., and Twum-Danso, N.A. (2007) Monitoring rare serious adverse events from a new treatment and testing for a difference from historical controls, *Clinical Trials*, **6**, 598–610.

Fontanarosa, P.B., Flanagin, A. and DeAngelis, C.D. (2005) Reporting conflicts of interest, financial aspects of research and role of sponsors in funded studies, *JAMA*, **294**, 110–111.

Frangakis, C.E. and Rubin, D.B. (2002) Principal stratification in causal inference, *Biometrics*, **58**, 21–29.

Fredriksson, T. and Pettersson, U. (1978) Severe psoriasis—oral therapy with a new retinoid, *Dermatologica*, **157**, 238–244.

Friedman, L. and DeMets, D. (1981) The data monitoring committee: how it operates and why, *IRB: Ethics and Human Research*, **3**, 6–8.

Friedman, M.A., Woodcock, J., Lumpkin, M.M. et al. (1999) The safety of newly approved medicines: do recent market removals mean there is a problem? *JAMA*, **281**, 1728–1734.

Gallo, P. (2006) Confidentiality and trial integrity issues for adaptive designs, *Drug Information Journal*, **40**, 445–449.

Gallo, P. and Krams, M. (2006) PhRMA working group on adaptive designs: introduction to the full white paper, *Drug Information Journal*, **40**, 421–423.

Gart, J. (1970) Point and interval estimation of the common odds ratio in the combination of 2×2 tables with fixed marginals, *Biometrika*, **57**, 471–475.

Gart, J. (1971) The comparison of proportions: a review of significance tests, confidence intervals and adjustments for stratification, *Review of the Indian Statistical Institute*, **39**, 148–169.

Gelman, A., Carlin, J.B., Stern, H.S. et al. (2004) *Bayesian Data Analysis*, Second Ed., Boca Raton: Chapman & Hall/CRC.

German Institute for Quality and Efficiency in Health Care. (2008) IQWiG Home Page http://www.iqwig.de/index.2.en.html.

Geyer, C.E., Forster, J., Lindquist, D. et al. (2006) Lapatinib plus capecitabine for HER2-positive advanced breast cancer, *New England Journal of Medicine*, **355**, 2733–2743.

Glaser, A. (2002) A set of quality criteria for statistical programming, *Drug Information Journal*, **36**, 565–570.

Glass, G.V. (1976) Primary, secondary and meta-analysis of research, *Educational Researcher*, **5**, 3–8.

Goldman, S.A. (2002) Adverse event reporting and standard medical terminologies: strengths and limitations, *Drug Information Journal*, **36**, 439–444.

Gray, R.J. (1988) A class of k-sample tests for comparing the cumulative incidence of a competing risk, *Annals of Statistics*, **16**, 1141–1154.

Grunkenmeier, G.L., Johnson, D.M., and Naftel, D.C. (1994) Sample size requirements for evaluating heart valves with constant risk events, *The Journal of Heart Valve Disease*, **3**, 53–58.

Haas, J.F. (2004) A problem-oriented approach to safety issues in drug development and beyond, *Drug Safety*, **27**, 555–567.

Haluska, F. and Aamdal, S. (2007) Personal communication.

Hamilton, M. (1960) A rating scale for depression, *Journal of Neurology, Neurosurgery and Psychiatry*, **23**, 56–62.

Harris, G. (2007) FDA issues strictest warning on diabetes drugs, *New York Times*, http://www.nytimes.com, accessed June 2007.

Havel, R.J., Hunninghake, D.B., Illingworth, D.R. et al. (1987) Lovastatin (mevinolin) in the treatment of heterozygous familial hypercholesterolemia—a multicenter study, *Annals of Internal Medicine*, **107**, 609–615.

Hemmings, R. and Day, S. (2004) Regulatory perspectives on data safety monitoring boards: protecting the integrity of data. *Drug Safety*, **27**, 1–6.

Herson J., Ognibene F.P., Peura, D.A. et al. (1992) The role of an independent data monitoring board in a clinical trial sponsored by a pharmaceutical firm, *Journal of Clinical Research and Pharmacoepidemiology*, **6**, 285–292.

Herson J. (1993) Data monitoring boards in the pharmaceutical industry, *Statistics in Medicine*, **12**, 555–561.

Herson J. (2008) Coordinating data monitoring committees and adaptive clinical trial designs, *Drug Information Journal*, **42**, 297–301.

Hibberd, P.L. and Weiner, D.L. (2004) Monitoring participant safety in phase I and II interventional trials: options and controversies, *Journal of Investigative Medicine*, **52**, 446–452.

Hollander, M. and Wolfe, D.A. (1999) *Nonparametric Statistical Methods*, Second Ed., New York: John Wiley.

Hosmer, D.W. and Lemeshow, S. (2000) *Applied Logistic Regression*, New York: Wiley.

International Conference on Harmonisation. (1994*) E2A: Clinical Safety Data Management: Definitions and Standards for Expedited Reporting*, http://www.ich.org/LOB/media/MEDIA436.pdf.

International Conference on Harmonisation. (1995) E3: *Structure and Content of Clinical Study Reports*, http://www.ich.org/LOB/media/MEDIA479.pdf.

International Conference on Harmonisation. (1996) *E6: Guideline for Good Clinical Practice*, http://www.ich.org/LOB/media/MEDIA482.pdf.

International Conference on Harmonisation. (1998) *E9: Statistical Principles for Clinical Trials*, http://www.ich.org/LOB/media/MEDIA485.pdf.

International Conference on Harmonisation. (2007) *MedDRA Data Retrieval and Presentation: Points to Consider*. http://www.ich.org/cache/html/607-272-1.html.

Ioannidis, J.P.A., Evans, S.J.W., Getzsche, P.C. et al. (2004) Better reporting of harms in randomized trials: an extension of the CONSORT statement, *Annals of Internal Medicine*, **141**, 781–788.

Ioannidis, J.P.A. and Lau, J. (2002) Improving safety reporting from randomized trials, *Drug Safety*, **25**, 77–84.

Ioannidis, J.P.A., Mulrow, C.D., and Goodman, S.N. (2006) Adverse events: the more you search, the more you find (Editorial), *Annals of Internal Medicine*, **144**, 298–300.

Johann-Liang, R., James, A.N., Behr, V.L. et al (2005) Reporting of deaths during pre-approval clinical trials for advanced HIV-infected populations, *Drug Safety*, **28**, 559–564.

Kahn, J. (2006) A nation of guinea pigs, *Wired*, 14.03, 277–283.

Kaplan, E.L. and Meier, P. (1958) Nonparametric estimation from incomplete observations, *Journal of the American Statistical Association*, **53**, 457–480.

Klein, J.P., Logan, B., Harhoff, M. et al. (2007) Analyzing survival curves at a fixed point in time, *Statistics in Medicine*, **26**, 4505–4519.

Kubler, J., Vonk, R., Belmel, S. et al. (2005) Adverse event analysis and MedDRA: business as usual or challenge? *Drug Information Journal*, **39**, 63–72.

Lachin, J.M. (2004) Conflicts of interest in data monitoring of industry versus publicly financed clinical trials, *Statistics in Medicine*, **23**, 1519–1521.

Lan, K.K.G., Hu, P., and Proschan, M.A. (2009) A conditional power approach to the evaluation of predictive power. Accepted, *Statistics in Biopharmaceutical Research.*

Lan, K.K.G. and Wittes, J. (1988) The B-value: a tool for monitoring data, *Biometrics*, **44**, 579–585.

Lasser, K.E., Allen, P.D., Woolhandler, S.J. et al. (2002) Timing of new black box warnings and withdrawals for prescription medications, *JAMA*, **287**, 2215–2220.

Lassere, M.N.D., Johnson, K.P., Boers, M. et al. (2005). Standardized assessment of adverse events in rheumatology clinical trials: summary of the OMERACT 7 drug safety module update, *Journal of Rheumatology*, **32**, 2037–2041.

Lee, Y.H., Ji, J.D., and Song, G.G. (2007) Adjusted indirect comparison of celecoxib verses rofecoxib on cardiovascular risk, *Pharmacology International*, **27**, 477–482.

Lenz, H.-J., Van Cutsem, E., Khambata-Ford, S. et al. (2006) Multicenter phase II and translational study of cetuximab in metastatic colorectal carcinoma refractory to irinotecan, oxaliplatin and fluoropyrimidines, *Journal of Clinical Oncology*, **24**, 4914–4921.

Lewis, P.A., O'Sullivan, M.M. and Rumfield, W.R. (1988) Significant changes in Ritchie scores, *Rheumatology*, **27**, 32–36.

Lin, Y.-L., Chern, H.-D., and Chu, M.-L. (2003) Hepatotoxicity in the review of clinical safety data, *Drug Information Journal*, **37**, 155–158.

Lipid Research Clinics Program. (1984a) The Lipid Research Clinics Coronary Prevention Trial results I, *Journal of the American Medical Association*, **251**, 351–364.

Lipid Research Clinics Program. (1984b) The Lipid Research Clinics Coronary Prevention Trial results II, *Journal of the American Medical Association*, **251**, 365–374.

Lurie, P. and Sasich, L.D. (1999) Safety of FDA-approved drugs, *JAMA*, **282**, 2297–2298.

Magin, P. and Smith, W. (2005) Isotretinoin, depression and suicide: a review of the evidence, *British Journal of General Practice*, **55**, 134–138.

Martin L.F., Booth F.V., Karlstadt R.G. et al. (1993) Continuous intravenous cimetidine decreased stress-related upper gastrointestinal hemorrhage without promoting pneumonia, *Critical Care Medicine*, **21**, 19–30.

MedDRA, (2007) Maintenance Support and Services Organization, http://www.meddramsso.com/mssoweb/index.htm.

Mehrotra, D.V. and Heyes, J. (2004) Use of the false discovery rate for evaluating clinical safety data, *Statistical Methods in Medical Research*, **13**, 227–238.

Melzack, R. and Torgerson, W.S. (1971) On the language of pain, *Anesthesiology*, **34**, 50–59.

Meyboom, R.H.B., Lindquist, M., and Egberts, A.C.G. (2000) An ABC of drug-related problems, *Drug Safety*, **22**, 414–423.

Mikuls, T.R. and Moreland, L.W. (2003) Benefit-risk assessment of infliximab in the treatment of rheumatoid arthritis, *Drug Safety*, **26**, 23–32.

Mundy, A. (2008) Drugs links to suicide risk draw concerns, *Wall Street Journal*, http://www.wsj.com, accessed July 9, 2008.

Ng, H.K.T. and Tang, M.-L. (2005) Testing the equality of two Poisson means using the rate ratio, *Statistics in Medicine*, **24**, 955–965.

Nissen, S.E., Wolski, K., and Topol, E.J. (2005) Effect of muraglitazar on death and major adverse cardiovascular events in patients with type 2 diabetes mellitus, *JAMA*, **294**, 2581–2586.

Nissen, S.E. and Wolski, K. (2007) Effect of rosiglitazone on the risk of myocardial infarction and death from cardiovascular causes, *New England Journal of Medicine*, **356**, 2457–2471.

O'Neill, R. (2007) Remarks made at Drug Information Association/FDA Statistics Forum, Bethesda, MD.

Pharmadhoc. (2007) *CTCAE Mapping Documents*, http://www.pharmadhoc.com/NCI%20CTC.htm.

Piantadosi, S. (2005) *Clinical Trials: A Methodological Perspective*, Second ed., Hoboken, NJ: John Wiley and Sons.

Pintilie, M. (2006) *Competing Risks: A Practical Perspective*, West Sussex, England: John Wiley and Sons.

Platonov, P. (2003) Clinical trials in Russia and Eastern Europe: recruitment and quality, *International Journal of Clinical Pharmacology and Therapeutics*, **41**, 277–280.

Pocock, S.J. (1983) *Clinical Trials: A Practical Approach*. New York: John Wiley & Sons.

Pounds, S. and Cheng, C. (2006) Robust estimation of the false discovery rate, *Bioinformatics*, **22**, 1979–1987.

Proschan, M.A., Lan, K.K.G., and Wittes, J.T. (2006) *Statistical Monitoring of Clinical Trials: A Unified Approach*, New York: Springer.

Rochester, G. (2008) Remarks made at Spring Statistical Meetings, Eastern North American Region, International Biometrics Society, Arlington, VA.

Rosenbaum, P.R. and Rubin, D.B. (1983) The central role of the propensity score in observational studies for causal effects, *Biometrika*, **70**, 41–55.

Rosner, B. (2006) *Fundamentals of Biostatistics*, Sixth Ed., Pacific Grove, CA: Duxbury Press.

Royall, R.M. (1997) *Statistical Evidence: A Likelihood Paradigm*, London: Chapman and Hall.

Royall, R.M. (2000) On the probability of observing misleading statistical evidence (with discussion), *Journal of the American Statistical Association*, **95**, 760–767.

Rush, A.J., Marangell, L.B., Sackheim, H.A. et al. (2005) Vagus nerve stimulation for treatment-resistant depression: a randomized controlled acute phase trial, *Biological Psychiatry*, **58**, 347–354.

Sackett, D.L. (1979) Bias in analytic research. *Journal of Chronic Disease*, **32**, 51–63.

SAS Institute. (2008) *SAS OnlineDoc 9.13, SAS/STAT User's Guide, Introduction to Survival Analysis Procedures*, http://support.sas.com/onlinedoc/913/docMainpage.jsp.

Scheiner, L.B. (1997) Learning versus confirming in clinical drug development. *Clinical Pharmacology and Therapeutics*, **61**, 275–291.

Schoenfeld, P. (1999) Gastrointestinal safety profile of meloxicam: a meta-analysis and systematic review of randomized controlled trials, *American Journal of Medicine*, **107**, 48S–54S.

Schuchman, M. (2007) Approving the vagus-nerve stimulator for depression, *New England Journal of Medicine*, **356**, 1604–1607.

Siegel, J.P., O'Neil, R.T., Temple, R. et al. (2004) Independence of the statistician who analyses unblinded data, *Statistics in Medicine*, **23**, 1527–1529.

Siu, L.L. and Rowinsky, E.K. (1998) A risk-benefit assessment of irinotecan in solid tumors, *Drug Safety*, **18**, 395–417.

Snapinn, S., Cook, T., Shapiro, D. et al. (2004) The role of the unblinded sponsor statistician, *Statistics in Medicine*, **23**, 1531–1533.

Strampel, W., Emkey, R., and Civitelli, R. (2007) Safety considerations with bisphosphonates for the treatment of osteoporosis, *Drug Safety*, **30**, 755–763.

Stump, D. (2007). Personal communication.

Sydes, M.R., Spiegelhalter, D.J., Altman, D.G. et al. (2004) Systematic qualitative review of the literature on data monitoring committees for randomized controlled trials, *Clinical Trials*, **1**, 60–79.

Temple, R. (2003). Remarks made at an FDA briefing on acceptable efficacy endpoints for oncology clinical trials, Rockville, MD.

Thijs, L., Celis, H., Kiowski, W. et al. (1995) Double-blind comparison of antihypertensive treatment with ramipril and piretanide, given alone or in combination, *Journal of Cardiovascular Pharmacology*, **26**, 33–38.

U.K. National Health Service. (2008) National Institute for Health and Clinical Excellence, http://www.nice.org.uk/.

U.S. Food and Drug Administration. (1997) *Expedited Safety Reporting Requirements for Human Drug and Biological Products,* http://www.fda.gov/cder/aers/fr07oc97.htm.

U.S. Food and Drug Administration. (2004) *Innovation Stagnation: Challenge and Opportunity on the Critical Path to New Medical Products*, http://www.fda.gov/oc/initiatives/criticalpath/whitepaper.html.

U.S. Food and Drug Administration. (2006) *Guidance for Clinical Trial Sponsors: Establishment and Operation of Clinical Trial Data Monitoring Committees,* http://www.fda.gov/cber/gdlns/clintrialdmc.pdf.

U.S. Food and Drug Administration. (2008) *MedWatch Home Page*, http://www.fda.gov/medwatch/.

U.S. National Cancer Institute. Cancer Therapy Evaluation Program, (2006). *Common Terminology Criteria for Adverse Events v3.0 (CTCAE),* http://ctep.cancer.gov/forms/CTCAEv3.pdf.

U.S. National Institute of Allergy and Infectious Disease, Division of Microbiology and Infectious Diseases (2007), *Adult Toxicity Table*, http://www3.niaid.nih.gov/NR/rdonlyres/401FA10C-8DF2-4162-8B21-2554129F4294/0/dmidadulttox.doc.

U.S. National Institutes of Health. (1998) *NIH policy for data and safety monitoring*, http://grants.nih.gov/grants/guide/notice-files/not98-084.html.

U.S. National Institutes of Health. (1999) *Guidance on reporting adverse events to Institutional Review Boards for NIH multicenter trials*, http://grants.nih.gov/grants/guide/notice-files/not99-107.html.

U.S. National Institutes of Health. (2000) *Further guidance on data and safety monitoring for phase I and II trials*, http://grants.nih.gov/grants/guide/notice-files/NOT-OD-00-038.html.

U.S. Supreme Court. (1964) *Jacobellis v. Ohio, 378 U.S. 184*, http://www.aegis.com/law/SCt/Decisions/1964/378US184.html.

University of Washington. (2007) *Adverse events: definition and grading*, http://www.crc.washington.edu/DataSafetyMonitoringPlans/Adverse.aspx.

Van Oers, N.H.J., Klasa, R., Marcus, R.E. et al. (2006) Rituximab maintenance improves clinical outcome of relapse resistant follicular non-Hodgkin's lymphoma in patients both with and without rituximab during induction: results of a prospective randomized phase III intergroup trial., *Blood*, **108**, 3295–3301.

Venulet, J. and Bankowski, Z. (1998) Harmonizing adverse drug reaction terminology: the role of the Council for International Organizations of Medical Sciences, *Drug Safety*, **19**, 165–172.

Wacker, B., Nagrani, T., Weinberg, J. et al. (2007) Correlation between development of rash and efficacy in patients treated with the epidermal growth factor receptor tyrosine kinase inhibitor erlotinib in two large phase III studies, *Clinical Cancer Research*, **13**, 3913–3921.

Weihrauch, T.R. and Kubler, J. (2002) Integrated summaries and meta-analyses in clinical drug development, *Drug Information Journal*, **36**, 127–133.

Wernicke, J.F., Faries, D., Milton, D. et al. (2005) Detecting treatment-emergent adverse events in clinical trials—a comparison of spontaneously reported and solicited collection methods, *Drug Safety*, **28**, 1057–1063.

Wernicke, J., Lledo, A., and Raskin, J. (2007) An evaluation of the cardiovascular safety profile of duloxetine: findings from 42 placebo-controlled studies, *Drug Safety*, **30**, 437–455.

White, C.A. (1998) A preliminary assessment of the impact of MedDRA on adverse event reports and product labeling, *Drug Information Journal*, **32**, 347–362.

Whitehead, A. (2002) *Meta-Analysis of Controlled Clinical Trials*, New York.

Wilson, J.F., Weale, M.F., Smith, A.C. et al. (2001) Population genetics structures of variable drug response, *Nature Genetics*, **29**, 265–269.

Wittes, J., Barrett-Connor, E., Braunwald, E. et al. (2007) Monitoring the randomized trials of the Women's Health Initiative: the experience of the data and safety monitoring board, *Clinical Trials*, **4**, 218–234.

Woodin, K.E. and Schneider, J.C. (2003). *The CRAs Guide to Monitoring Clinical Research*, Boston: Thompson CenterWatch.

Writing Group for the Women's Health Initiative Investigators. (2002) Risks and benefits of estrogen plus progestin in health postmenopausal women—principal results from the Women's Health Initiative randomized controlled trial, *JAMA*, **288**, 321–333.

Wysowski, D.K. and Swartz, L. (2005) Adverse drug event surveillance and drug withdrawals in the United States, 1969–2002, *Archives of Internal Medicine*, **165**, 1363–1369.

Ziemssen, T., Neuhaus, O., and Hohlfeld, R. (2001) Risk-benefit assessment of glatiramir acetate in multiple sclerosis, *Drug Safety*, **24**, 979–990.

Index

E

N

O

P

Q

R

S

T

U

V

W

Z

Statistics

In the world of drug development, clinical issues and statistical issues cannot be separated. All are scientific and all use applied logic. However, safety monitoring in pharmaceutical industry clinical trials through data monitoring committees (DMCs) is both an art and a science, requiring the use of sound judgment as well as sound scientific method.

In **Data and Safety Monitoring Committees in Clinical Trials**, Jay Herson applies his years of experience serving on and providing statistical support to DMCs. He reviews the current state of DMCs and the best practices that have evolved using the same lively approach he employed when writing *Herson's Handout* for the ALA newsletter *Under the Curve*, 1991–2004. Defining the stewardship role and inner workings of DMCs, Dr. Herson —

- Describes a DMC's interactions with sponsors, data analysis centers, institutional review boards, and regulatory agencies
- Examines the biases and pitfalls in analyzing safety data, as well as other clinical issues including the distinctions among adverse events, serious adverse events, and severe adverse events
- Presents those statistical methods useful for DMCs, illustrated with data from actual clinical trials
- Demonstrates how physicians think differently than statisticians about safety data, and explains why both views are needed

With regulatory agencies worldwide now facing considerable challenges, it is important that those involved with trials review the direction and effectiveness of DMCs. Providing a perspective that few can match, Dr. Herson fully explains the types of decisions DMCs are called upon to make as well as the environment in which they are made, taking into consideration the cultural, political, and clinical issues that require the use of good sense along with good science.

CRC Press
Taylor & Francis Group
an **informa** business
www.crcpress.com
6000 Broken Sound Parkway, NW
Suite 300, Boca Raton, FL 33487
270 Madison Avenue
New York, NY 10016
2 Park Square, Milton Park
Abingdon, Oxon OX14 4RN, UK

C7037
ISBN: 978-1-4200-7037-8
90000
9 781420 070378