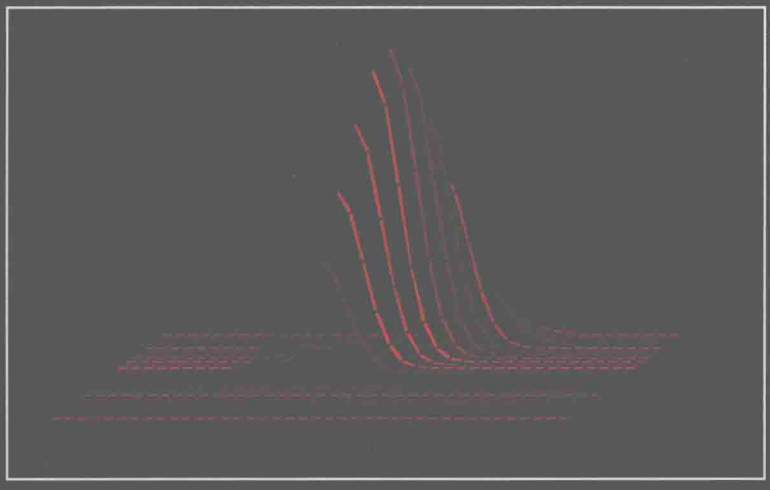
Lemuel A. Moyé

Statistical Monitoring of Clinical Trials

Fundamentals for Investigators



Springer

Lemuel A. Moyé

Statistical Monitoring of Clinical Trials

Fundamentals for Investigators

With 69 Figures

Springer

Lemuel A. Moyé
University of Texas
Health Science Center at Houston
School of Public Health
Houston, TX 77030
USA
moyelaptop@msn.com

Library of Congress Control Number: 2005931408

ISBN-10: 0-387-27781-1 e-ISBN 0-387-27782-X
ISBN-13: 978-0387-277813

Printed on acid-free paper.

Printed in the United States of America. (MVY)

9 8 7 6 5 4 3 2 1

springeronline.com

Statistical Monitoring of Clinical Trials

To Dixie and the DELTs

Preface

A preface is an opportunity for you and me to share an amiable conversation before the serious work starts. If you give me a moment, I will share with you my motivations for writing an introductory text about the statistical monitoring of clinical trials, a staple of modern research efforts in heatlhcare.

I am pleased to have been involved in clinical research for eighteen years. Many of my efforts focused on preparations for and presentations to Data Monitoring Committees (DMCs), each of which was tasked with overseeing the conduct of a particular clinical study. During these activities, I have spoken with many clinicians about the epidemiology and biostatistical foundation of this mode of clinical research.

In my experience, nothing confuses a DMC member as do these so-called "stopping rules" for monitoring the conduct of a healthcare research study. The idea of prematurely ending a study makes intuitive sense to the clinical members of the committee. The rules themselves with their arcane terminology are the problem. Descriptions of "group sequential procedures" and "stochastic curtailment" provide no useful handholds for the clinician working to understand this slippery but essential subject. The fact that neither medical school nor residency curricula discuss any of the details of these procedures is one possible explanation for the continued lack of understanding among clinicians. In general, the non-statistical members of newly conceived DMCs in the 21st century are just as confused about statistical monitoring guidelines as were their clinical predecessors who sat on DMCs in the 1980s.

A major reason for this continued confusion is that clinical investigators, although blessed with the motivation to do research, commonly do not have strong mathematical backgrounds. Although many have worked hard to develop the basic understanding of epidemiology and biostatistics necessary to be an effective investigator, the underlying mathematical details of commonly used monitoring procedures as frequently presented remain beyond the scope of their training.

Of course, the statistical literature has much to say on the subject of monitoring rules in clinical research. Beginning with the manuscripts of Armitage and Wald in the 1940s, the statistical treatment of this topic slowly expanded until the late 1970s, when it exploded. The recognition of the importance of the monitoring of clinical research, in concert with the complexity of the underlying mathematics has attracted the best and the brightest of biostatisticians. Their devotion to the study of the underlying mathematical structure of monitoring procedures has resulted in a body of knowledge that is both evolutionary and illuminating. However, because it tends to be scripted in the technical and exclusionary language of advanced mathematics, the writing tends to enlighten only the sophisticated analyst.

The text by Christopher Jennison and Bruce Turnbull [1] is a fine example of a comprehensive treatment of a difficult statistical subject.

The technical writing style that has been implemented in the field of "interim monitoring" should come as no surprise. However, work in this area, propelled forward by the strong rowing of capable statistical theorists, can leave the clinical investigator behind in its wake. Required to apply complex processes that they do not understand, the clinical investigator commonly finds little introductory material available. In addition, clinical researchers with no quantitative background have difficulty communicating with biostatisticians or experienced trial methodologists who have much experience but little time to explain these issues to their inexperienced colleagues. Thus, investigators who wish to learn about these mathematical procedures are hard pressed to identify readily understandable source material.

The purpose of this text is to fill that gap. If you know nothing about monitoring guidelines in clinical trials, then this book is for you.

I have chosen to begin this book with a brief history of monitoring rules in clinical research. Although this is the first chapter in this book, it needn't be the first chapter that you read. Being nontechnical, it might be most useful to view its contents as a pleasant oasis in a desert of more complicated discussion. Its considerations of the interactions between scientists serves to convey something about the people who were involved in these important historical efforts. The observation that the epidemiologist Bradford Hill suffered from tuberculosis years before he helped design an early clinical trial to study this disease may be a mere curiosity to some; to others it helps to explain his intellectual fortitude in working with skeptical clinicians.

For the same reason, I have broken up some of the technical arguments that appear in later chapters with an occasional vignette. As my students frequently remind me, it is best to have a joke close by when discussing anything mathematical.

I must confess that this is not a book about the operation of DMCs. That material has been very nicely developed in *Data Monitoring Committees in Clinical Trials: A Practical Perspective* by Susan Ellenberg, Thomas Fleming, and David DeMets (John Wiley & Sons, Ltd., West Sussex, 2002). Their text is very broad in scope, focusing on the DMCs evolution and contemporary operation. Our focus here is on statistical monitoring procedures that these DMCs devise and utilize, not on the DMCs themselves.

One final note. An important segment of the current clinical investigator population is comprised of women. Therefore, I have alternated the use of gender in the hypothetical illustrations offered by this text. Although this is the most illustrative and the least exclusionary approach, it does require mental alacrity on your part as the genders change from example to example.

Lemuel A. Moyé
The University of Texas
School of Public Health
Houston, Texas
July, 2005

References

1. Jennison C, Turnbull BW. (2000). *Group Sequential Methods with Applications to Clinical Trials*. New York. Chapman & Hall/CRC.

Acknowledgments

My understanding of monitoring procedures was not generated by reading complicated treatises, but instead was forged in vibrant and sometimes fiery discussions during Data Monitoring Committee meetings. The distinguished members serving on these panels deserve much of the credit for the educational material this book contains. Special thanks goes to the members of the University of Texas School of Public Health Coordinating Center for Clinical Trials, and to its senior members both past and present. Mort Hawkins ScD, Barry Davis, MD, PhD, and Robert Hardy, PhD. served as the guide rails for me, keeping me on the right track when I tended to veer too far in one direction or the other.

I also acknowledge the members of the Houston SPOTRIAS team. In this effort, we investigators have faced numerous challenges in working to devise useful monitoring guidelines for small pilot research projects, and in doing so, have created through debate and discussion many good ideas. Special thanks to Dr. James Grotta, Andre Alexandrov and Louis Morgenstern for their roles in these conversations.

John Kimmel, the reviewing editor, and several chapter reviewers have provided good questions, challenges, suggestions, and additional references. Their help has improved this book's structure.

Nidhi Rohatgi, my graduate teaching assistant, provided invaluable assistance in clarifying the text and in copyediting. Nidhi was especially helpful in pointing out the need for me to untangle what had become relatively impenetrable tracts of some chapters. Dr. Claudia Pedroza provided fine support for the discussion of Bayesian monitoring guidelines.

Finally, my dearest thanks go to Dixie, my wife, on whose personality, character, love, and common sense I have come to rely, and to Flora and Bella Ardon, whose continued emotional and spiritual growth reveals anew to me each day that, through God, all things are possible.

Lemuel A. Moyé, MD, PhD
University of Texas
School of Public Health
Houston, Texas, USA
July, 2005

Introduction

Statistical monitoring procedures are the body of computations that aid clinical investigators in determining if a research program should be suspended prematurely. Specifically, these guidelines are used to guide the complex decision to end a clinical study if the investigation is very likely to produce either (1) an early positive benefit, (2) an early indication of harm, or (3) a neutral effect at the time the study is scheduled to end (expressed as stopping for "futility"). Research scientists and members of clinical trial oversight committees rely upon these procedures, colloquially expressed as "stopping rules", but more correcting described as "monitoring guidelines".

Although clinical investigators accept the application of statistical and epidemiologic principles in clinical research, the procedures used to terminate clinical studies often appear opaque to the statistically naïve investigator. Nevertheless, these guidelines have become ubiquitous in healthcare research. In 1998, the Office of the Inspector General of the Department of Health and Human Services mandated that the Food and Drug Administration (FDA) and the National Institutes of Health (NIH) develop such procedures and standards for U.S. trials. In response, the NIH has generated policies to require safety monitoring plans for all phase III NIH-funded studies, and the FDA has issued a draft guidance document on the establishment and operation of the committees that perform such monitoring. In addition, the Institutional Review Boards (IRBs) that govern the ethical conduct of clinical investigation at many research centers developed their own sets of instructions for the application of oversight procedures. These monitoring responsibilities reside in the Data Monitoring Committees (DMCs) of the individual clinical research projects.

This new requisite for formal statistical monitoring of clinical research places clinical investigators in a dilemma. As researchers in a study, they have to satisfy the monitoring requirements of their institutional review board. Alternatively, if they are members of a DMC, then their input into the discussions that calibrate the statistical monitoring device of the study is required. However, these investigators are commonly ill equipped to deal with the issues of modern statistical monitoring of clinical trials. Thus, they are unable to fruitfully engage in the discussion, development, or defense of the use of these tools.

Well-motivated, but statistically unsophisticated clinical investigators can learn the correct use and interpretation of these monitoring procedures when provided with a learning tool that informs them in clear language. This tool would allow them to steadily increase their knowledge of, experience with, and intuition about these procedures. *Statistical Monitoring of Clinical Trials: Fundamentals for Investigators* is this tool. Specifically, it provides the discussion of these statistical devices that clinical investigators need, representing a user-friendly introduction to

monitoring procedures for these scientists. These essential statistical considerations are rarely taught in introductory biostatistics or medical statistics classes.

Chapter One of *Statistical Monitoring of Clinical Trials: Fundamentals for Investigators* provides an overview of the evolution of monitoring procedures in clinical research. Randomized, blinded controlled clinical trials, available for only sixty years, are a relatively new tool in clinical investigation, and remain controversial. The ethical concerns raised by this investigational methodology have called for the interim monitoring of these studies. This demand in turn has generated a relatively new application for Brownian motion, one completely unforeseen by its progenitors, including Albert Einstein.

Chapter Two provides a review of the basic statistical thought process required in clinical research and directly applicable to interim monitoring. The set of circumstances that permit one to generalize the results from a single small sample to a population of thousands or millions of subjects has direct bearing on the successful application of statistical monitoring of clinical trials. These situations and their limitations are discussed in detail. In addition, the foundation principles of statistical hypothesis testing, confidence intervals, and the Bayes approach are each described.

Chapter Three develops the elementary principles of probability that are required to understand the principles behind the interim review of clinical research results. The differences between subjective and objective probability are discussed, and the roles of each in the statistical monitoring of clinical trials are explained. In addition, the concept of probability as an area under a curve is illuminated, with special emphasis given to the normal distribution. Finally, elementary examples of the use of probability for the early termination of a clinical research effort are provided. Chapters Two and Three provide the foundation for the rest of the text.

Chapter Four addresses the need for monitoring procedures in clinical research. This chapter lays out for the clinical scientist the problems that arise when one attempts to use traditional hypothesis testing procedures to draw conclusions about a clinical study's interim results. It provides, through the use of discussion and examples, the elaboration the clinical scientist needs in order to develop insight into the basic behavior of statistical monitoring tools. Investigators have become familiar with the idea of a test statistic's location (i.e., whether the test statistic is greater than 1.96). In this chapter, that notion is supplemented with the observation that a test statistic follows a particular path to arrive at its current location. An examination of that path's properties reveals new information that can provide accurate predictions of the test statistic's location in the future. This concept is new to most clinical investigators, and is elaborated in detail without heavy reliance on mathematics. It is here that the link between Brownian motion and clinical monitoring procedures is motivated.

Capitalizing on the insight provided in Chapter Four, Chapter Five introduces the basic group sequential approach of Pocock and O'Brien–Fleming, followed by discussions of the Haybittle–Peto and Lan–DeMets derivatives. The triangular designs popularized by Whitehead are briefly discussed. Chapter Six develops conditional power in a way that illuminates the circumstances in which a clinical trial may be stopped early for a beneficial finding based on a "look forward" approach.

Chapter Seven describes the use of monitoring procedures to identify harmful effects of the tested intervention. This is a natural introduction to the current use of asymmetric monitoring procedures. In addition, the problem of deciding to discontinue a study because of an unanticipated finding in one of several safety measures is developed. The many unexpected safety considerations that can arise during the study's execution amplify the importance of this issue. This chapter also introduces the notion of stopping a clinical trial early due to "futility".

Chapter Eight provides an introduction to the use of monitoring procedures using the Bayes paradigm. Each chapter ends with a relevant problem set.

This book can serve as a reference text for clinical scientists at all levels of training, being especially useful for healthcare graduate students and junior physician-scientists. Its readers require basic college algebra, plus one course in healthcare statistics. Its contents are of interest to students attending medical schools, graduate schools with an emphasis in healthcare research, and schools of public health. In addition, the contents of *Statistical Monitoring of Clinical Trials: Fundamentals for Investigators* are applicable to workers in health departments, private institutes, and government regulatory agencies. This book is also useful for judges who, not uncommonly, have to learn about the ethical conduct (and, therefore, the ethical monitoring) of clinical research efforts.

This text's incorporation of background material as well as in-depth discussion requires some guidance for its optimal use. There are several sections in Chapters 5, 6, and 7 which have a "*" in their title, signifying that the material is more challenging for students with a weak background in probability. In addition, the appendices, providing some in-depth mathematical development, can also appear formidable to a student with one background course in statistics.

Therefore, this book may be successfully used as the basis for a basic, introductory course on monitoring rules in clinical trials by focusing on Chapters 1 through Chapter 7, ignoring (1) all of the starred sections in Chapters 5, 6, and 7 and (2) the contents of Appendices A through D. However, those with a stronger mathematics background, after reviewing the historical introduction, can move directly to Chapter 4 and proceed through Chapter 8, covering the details of Appendices A through E as needed.

One caveat. Healthcare researchers regardless of their level of mathematical sophistication, should spend some time in Chapter Two, which discusses the statistical reasoning process in medicine. The experience of the author is that, without this review, many researchers unfortunately use statistical monitoring procedures as a tool to identify the "smallest p-value the quickest way" leading to important setbacks in the development of both research programs and research careers.

Contents