Bayesian Adaptive Methods for Clinical Trials

Scott M. Berry, Bradley P. Carlin, J. Jack Lee, and Peter Müller

To

Our families

Contents

Foreword

It's traditional to get a foreword written by an *éminence grise*, generally an aging researcher who has seen better days. I can provide plenty of *grise* although I am possibly a bit short on *éminence*. Perhaps I best qualify through sheer long-service in trying to promote Bayesian clinical trials, having started my small contribution to this epic effort nearly 30 years ago with Laurence Freedman, eliciting prior opinions from oncologists about the plausible benefits of new cancer therapies.

This fine book represents the most recent and exciting developments in this area, and gives ample justification for the power and elegance of Bayesian trial design and analysis. But it is still a struggle to get these ideas accepted. Why is this? I can think of four main reasons: ideological, bureaucratic, practical and pragmatic.

By *ideological*, I mean the challenge facing the "new" idea of using probability theory to express our uncertainty about a parameter or existing state of the world – our epistemic uncertainty. Of course "new" is ironic, given it is nearly 250 years since Bayes formalized the idea, but the idea is still unfamiliar and disturbing to those brought up on classical ideas of probability as long-run frequency. One can only sympathize with all that effort to master the correct definition of a p-value and a confidence interval, only to be told that the intuitive meanings can be right after all.

I really enjoy introducing students to this beautiful idea, but tend to leave Bayes' theorem to subsequent lectures. In fact I sometimes feel the role of Bayes' theorem in Bayesian analysis is overemphasized: the crucial element is being willing to put a distribution over a parameter, and it is not always necessary even to mention the "B-word." Natural examples include models for informative dropout in clinical trials, and the size of possible biases in historical studies: in these situations there may be no information in the data about the parameter, and so Bayes' theorem is not used.

But of course there are *bureaucratic* obstacles: as the authors of this book make clear, regulatory agencies perform a gate-keeping role where the Neyman-Pearson framework of decision-making without a loss function still has merits. Although the posterior distribution tells us what it is reasonable to believe given the evidence in a specific study, the regulators do need to

consider a continuous sequence of drug approval decisions. So quantifying Type I and Type II error can still be a valuable element of trial design, and one that is excellently covered in this book.

Then there are *practical* problems: can we actually do the analysis, or is the mathematics too tricky and there's no software to help us along? The authors have done a great job in discussing computation and providing software, but I am sure would still admit that there's some way to go before all these wonderful techniques are easily available to the average trial designer. But it will happen.

Finally, the crucial *pragmatic* test. Do these techniques help us do things we could not do before? This has been the factor that has led to increasingly widespread penetration of Bayesian methods into subject domains over the last 20 years or so: people can fit models and make inferences that were previously impossible or very cumbersome. And this is where this book wins hands down, since adaptive trials are so natural, ethical and efficient, that everyone wants to do them.

This book, based on the many years of cumulative experience of the authors, manages to deal with all these difficulties. Adaptive studies are a perfect application for a Bayesian approach, and I am confident that this book will be a major contribution to the science and practice of clinical trials.

David J. Spiegelhalter
MRC Biostatistics Unit and University of Cambridge
April 2010

Preface

As has been well discussed, the explosion of interest in Bayesian methods over the last 10 to 20 years has been the result of the convergence of modern computing power and efficient Markov chain Monte Carlo (MCMC) algorithms for sampling from and summarizing posterior distributions. Practitioners trained in traditional, frequentist statistical methods appear to have been drawn to Bayesian approaches for three reasons. One is that Bayesian approaches implemented with the majority of their informative content coming from the current data, and not any external prior information, typically have good frequentist properties (e.g., low mean squared error in repeated use). Second, these methods as now readily implemented in `WinBUGS` and other MCMC-driven software packages now offer the simplest approach to hierarchical (random effects) modeling, as routinely needed in longitudinal, frailty, spatial, time series, and a wide variety of other settings featuring interdependent data. Third, practitioners are attracted by the greater flexibility and adaptivity of the Bayesian approach, which permits stopping for efficacy, toxicity, and futility, as well as facilitates a straightforward solution to a great many other specialized problems such as dose-finding, adaptive randomization, equivalence testing, and others we shall describe.

This book presents the Bayesian adaptive approach to the design and analysis of clinical trials. The ethics and efficiency of such trials can benefit from Bayesian thinking; indeed the Food and Drug Administration (FDA) Center for Devices and Radiological Health (CDRH) has been encouraging this through its document *Guidance for the Use of Bayesian Statistics*; see `www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm071072.htm`. The FDA Center for Drug Evaluation and Research (CDER) and Center for Biologics Evaluation and Research (CBER) has issued its own *Guidance for Industry: Adaptive Design Clinical Trials for Drugs and Biologics*: `www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM201790.pdf`. This document also mentions Bayes, albeit far less prominently. The recent series of winter Bayesian biostatistics conferences at the University

of Texas M.D. Anderson Cancer Center in Houston are also testament to the growing role Bayesian thinking plays in this field.

The outline of the book is as follows. In Chapter 1 we summarize the current state of clinical trial design and analysis, present the main ideas behind the Bayesian alternative, and describe the potential benefits of such an alternative. We also describe what we mean by the word "adaptive" in the book's title. Chapter 2 then gives an overview of the basic Bayesian methodological and computational tools one needs to get started as a Bayesian clinical trialist. While this whirlwind tour is not a substitute for a full course in Bayesian methods (as from Gelman et al., 2004, or Carlin and Louis, 2009), it should enable those with a basic understanding of classical statistics to get "up and running" on the material. This chapter also includes overviews of hierarchical modeling (with special emphasis on its role in Bayesian metaanalysis) and the basics of Bayesian clinical trial design and analysis. The idea here is to establish the basic principles that will be expanded and made phase- and endpoint-specific in subsequent chapters.

The next two chapters of the book (Chapters 3–4) follow standard clinical trials practice by giving Bayesian tools useful in "early" and "middle" phase clinical trials, roughly corresponding to phases I and II of the U.S. drug regulatory process, respectively. While our own professional affiliations have led us to focus primarily on oncology trials, the techniques we describe are readily adapted to other disease areas. We also place primary emphasis on "partially Bayesian" designs that concentrate on probability calculations utilizing prior information and Bayesian updating while still maintaining good frequentist properties (power and Type I error). An exception to this general rule is Section 4.6, where we discuss "fully Bayesian" designs that incorporate a utility function (and often more informative priors) within a more formal decision-theoretic framework. Chapter 4 also contains brief reviews of two recent trials utilizing Bayesian adaptive designs, BATTLE and I-SPY 2.

Chapter 5 deals with late (phase III) studies, an important area and the one of potentially greatest interest to statisticians seeking final regulatory approval for their compounds. Here we emphasize modern adaptive methods, seamless phase II–III trials for maximizing information usage and minimizing trial duration, and describe in detail a case study of a recently approved medical device. Finally, Chapter 6 deals with several important special topics that fit into various phases of the process, including the use of historical data, equivalence studies, multiplicity and multiple comparisons, and the related problem of subgroup analysis. The historical data material is particularly relevant for trials of medical devices, where large historical databases often exist, and where the product being evaluated (say, a cardiac pacemaker) is evolving slowly enough over time that worries about the exchangeability of the historical and current data are relatively low.

Since this is not a "textbook" per se, we do not include homework prob-

lems at the end of every chapter. Rather, we view this book as a handbook enabling those engaged in clinical trials research to update and expand their toolkit of available techniques, so that Bayesian methods may be used when appropriate. See `http://www.biostat.umn.edu/~brad/data3.html` and `http://biostatistics.mdanderson.org/SoftwareDownload/` on the web for many of our datasets, software programs, and other supporting information. The final sections of Chapters 2–6 link to these software sites and provide programming notes on the `R` and `WinBUGS` code we recommend.

We owe a debt of gratitude to those who helped in our writing process. In particular, the second author is very grateful to Prof. Donald Berry and the Division of Quantitative Sciences at the University of Texas M.D. Anderson Cancer Center for allowing him to spend his fall 2008 sabbatic time in the same U.S. state as the other three authors. Key staff members worthy of special mention are Martha Belmares and the incomparable Lydia Davis. Sections 1.1, 1.2, 1.4, and 2.4 are based on Prof. Berry's previous work in their respective areas. Indeed, many sections of the book owe much to the hard work of our research colleagues, including Lee Ann Chastain, Nan Chen, Jason Connor, Laura Hatfield, Brian Hobbs, Haijun Ma, Ashish Sanil, and Amy Xia. We also thank the 2010 spring semester "Topics in Clinical Trials" class at Rice University and the University of Texas Graduate School of Biomedical Sciences," taught by the third author, for commenting on the text and testing the supporting software. Rob Calver and David Grubbs at Chapman and Hall/CRC/Taylor & Francis Group were pillars of strength and patience, as usual. Finally, we thank our families, whose ongoing love and support made all of this possible.

Scott M. Berry College Station, Texas
Bradley P. Carlin Minneapolis, Minnesota
J. Jack Lee Houston, Texas
Peter Müller Houston, Texas

March 2010

Index