1

PRELIMINARIES

1.1 INTRODUCTION

The best time to contemplate the quality of evidence from a clinical trial is before it begins. High-quality evidence about the effects of a new treatment is a consequence of good study design and execution, which themselves are the results of careful planning. This book presents elements of clinical trial methodology that are needed in planning, designing, conducting, analyzing, and assessing clinical trials with the goal of improving the evidence derived from these important studies. The topics discussed include subjects relevant to early planning and design—some that find general agreement among methodologists and others that are contentious. It is unlikely that a reader experienced with clinical trials will agree with all that I say or emphasize, but my perspective should be mainstream, internally consistent, and useful for learning.

Much of what is written about clinical trials tells us what to do. As a result, many investigators, sponsors, and regulators become proficient in standard practice, perhaps to the point that they view it as restrictive. Not as much of the literature explains why we do what is recommended. Knowing only what to do creates patterned thinking, whereas knowing why we do it allows appropriate creative exceptions and coping with atypical circumstances. I try throughout this book to emphasize the rationale for what we do in clinical trials, and specifically discourage unhelpful patterned thinking.

My further purpose is to challenge the dogma and traditions to which many young investigators have been exposed by their older colleagues. I will not stereotype and validate impressions gathered from the very imperfect culture of clinical investigation that exists nearly everywhere. In recent years especially, I have noticed some aggressive ignorance regarding the scientific principles underlying clinical trials. The goal of this book is to teach foundational principles and their sometimes complicated implications so
that the sublime, ordinary, and absurd in clinical trials becomes more clear. As a general rule, I will discard cherished perspectives if they lack a strong scientific basis.

This book is not an introduction to clinical trials and has evolved to become more technical and lengthy as it has been updated through the years. Even so, I am aware from colleagues that portions of it have found their way into didactics for beginning students of clinical trials. My intent for this book is to be part of a one quarter or semester structured postgraduate course for an audience with quantitative skills and therapeutic focus. I take for granted that the audience has already completed a formal didactic introduction to clinical trials. I also expect that the audience will have completed a basic postgraduate series in biostatistics. The first edition of this book evolved over a dozen years by merging a course in experimental design with one on clinical trials. The second edition was the result of seven additional years of teaching and concomitant changes in the field. The third edition incorporates changes and needs that I perceive in the field over the past 8–10 years, and responds partly to my experiences teaching diverse clinical investigators at a second major medical school and academic medical center.

1.2 AUDIENCES

The intended teaching audience includes both medical and statistical professionals early in their clinical research careers. Each component of the audience presents a certain dilemma. The book assumes a working knowledge of basic biostatistics. It is not really possible to get beyond the introductory ideas of clinical trials without statistics. It is also helpful if the reader understands some more advanced statistical concepts, including ideas like methods of inference, error control, lifetables, survival models, and likelihoods. Clinicians often lack this knowledge. However, modern clinical researchers are seeking the required quantitative background through formal training in clinical investigation methods or experimental therapeutics.

Biostatistics professionals require no clinical knowledge to understand the concepts in this book. Even so, fundamental principles of clinical research are helpful, including some basic human biology, research ethics, therapeutic developmental paradigms, clinical terminology and ideas of clinical pharmacology. Eventually, for the clinical trials biostatistician, immersion in a disease-oriented discipline will be vital to make the most substantive collaborative contributions as well as to identify and solve relevant methodological problems.

Uneven discussion (basic to technically complex) results from this mixed target audience and perspectives, as well as from the very nature of clinical trials. The classes that I teach using this book typically contain a mixture of biostatistics graduate students, medical doctors in specialty or subspecialty training (especially working toward a degree in clinical investigation), and other health professionals training to be sophisticated managers or consumers of clinical trials. For such an audience, the goal is to provide breadth and to write so as not to be misunderstood. This book should be supplemented with lecture and discussion, and possibly a computer lab. The reader who does not have an opportunity for formal classroom dialogue will need to explore the references more extensively. Exercises and discussion questions are provided at the end of each chapter. Most are intentionally made open-ended, with a suggestion that the student answer them in the form of a one- or two-page memorandum, as though providing an expert opinion to less-experienced investigators.
In short, this book targets clinical trialists, who are not so simple to define. Operationally, a clinical trialist is someone whose career is focused on the science of trials. A trialist uses investigational methods across disciplinary boundaries, compared to a specialist who might perform some trials in a single domain. Being true interdisciplinary experts, trialists can be derived from a number of sources: (i) quantitative or biostatistical, (ii) administrative or managerial, (iii) clinical, or (iv) ethics. Students can effectively approach the subject from any of these perspectives. Eventually however, a mature trialist will be conversant with important details from all the foundational fields.

It is common today for rigorous trialists to be strongly statistical. This is because of the fairly rapid recent pace of methods for clinical trials coming from that field, and because statistics pertains to all of the disciplines in which trials are conducted (or all of science for that matter). However, the discussion in this book does not neglect the other viewpoints that are also essential to understanding trials. Many examples herein relate to cancer because that is the primary field in which I work, but the concepts will generalize to other areas. Studying trials in different clinical disciplines is the best device for understanding principles. Unfortunately, the structure of many institutions and collaborations inhibits this.

Scientists who specialize in clinical trials are frequently dubbed “statisticians” by their clinical colleagues. I will sometimes use that term with the following warning regarding rigor: statistics is an old and broad profession. There is not a one-to-one correspondence between statisticians or biostatisticians and knowledge of clinical trials. However, trial methodologists, whether statisticians or not, are likely to know a lot about biostatistics and will be accustomed to working with statistical experts. Many trial methodologists are not statisticians at all, but evolve from epidemiologists or clinicians with a strongly quantitative orientation, as indicated above. My personal experience is that there are many excellent clinical trialists whom statisticians would label “physicians.” The stereotyping is not important. The essential idea is that the subject has many doorways.

In recent years, many biostatistics professionals have gravitated toward bioinformatics, computational biology, or high-dimensional data. I shall not take the space here to define these terms or distinguish them from research informatics or (bio)medical informatics. It is just worth pointing out that many biostatisticians have no experience with clinical trials. Technologies are of high value, but each wax and wanes with limited impact on methods of comparison.

The principles of valid comparisons in medicine have evolved over hundreds of years, and especially in the recent century. These principles have not been altered by the phenomenal ongoing scientific or technological advances during that period of time. Revolutions such as germ theory, drug development, anesthesia and surgery, immunology, genomic science, biomedical imaging, nutrition, computerization, and all the others of history have not diminished the foundational need of medicine to assay treatments fairly. In fact, as these revolutions integrated themselves into scientific medicine, the need for therapeutic comparison expands and the principles for doing so remain constant.

1.3 SCOPE

I have made an effort to delineate and emphasize principles common to all types of trials: translational, developmental, safety, comparative, and large-scale studies. This follows
from a belief that it is more helpful to learn about the similarities among trials rather than differences. However, it is unavoidable that distinctions must be made and the discussion tailored to specific types of studies. I have tried to keep such distinctions, which are often artificial, to a minimum. Various clinical contexts also treat trials differently, a topic discussed briefly in Chapter 4.

There are many important aspects of clinical trials not covered here in any detail. These include administration, funding, conduct, quality control, and the considerable infrastructure necessary to conduct trials. These topics might be described as the technology of trials, whereas my intent is to focus on the science of trials. Technology is vitally important, but falls outside the scope of this book. Fortunately, there are excellent sources for this material.

No book or course can substitute for regular interaction with a trial methodologist during both the planning and analysis of a clinical investigation. Passive reliance on such consultations is unwise, but true collaborations between clinicians and trialists will result when both grasp the relevant body of knowledge. Although many clinicians think of bringing their final data to a statistician, a collaboration will be most valuable during the design phase of a study when an experienced trialist may prevent serious methodologic errors, create efficiencies, or help avoid mistakes of inference.

The wide availability of computers is a strong benefit for clinical researchers, but presents some hidden dangers. Although computers facilitate efficient, accurate, and timely keeping of data, modern software also permits or encourages researchers to produce “statistical” reports without much attention to study design and without fully understanding assumptions, methods, limitations, and pitfalls of the procedures being employed. Sometimes a person who knows how to run procedure-oriented packages on computerized data is called the “statistician,” even though he or she might be a novice at the basic theory underlying the analyses. It then becomes possible to produce a final report of a study without the clinical investigator understanding the study execution, the limitations of analysis, and without the analyst being conversant with the data. What a weak chain that is.

An additional warning about the limits of technology is necessary. There is no computerized summary of clinical data that can ensure its correctness. Inspection via a computer-generated printout of data values by a knowledgeable investigator can help re-ensure us that gross errors are unlikely. But once the data are collapsed, re-coded, tabulated, transformed, summarized, or otherwise combined with each other by the computer, errors are likely to be hidden. Orderly looking summaries do not ensure correctness of the underlying data.

The ideas in this book are intended to assist reliability, not by being old-fashioned but by being rigorous. Good design inhibits errors by involving a statistical expert in the study as a collaborator from the beginning. Most aspects of the study will improve as a result, including reliability, resource utilization, quality assurance, precision, and the scope of inference. Good design can also simplify analyses by reducing bias and variability and removing the influence of irrelevant factors. In this way number crunching becomes less important than sound statistical reasoning.

The student of clinical trials should also understand that the field is growing and changing in response to both biological and statistical developments. A picture of good methodology today may be inadequate in the near future. Designing trials to use biomarkers and genomics data effectively is one area of evolution. But change is probably even more true of analytic methods than design, where the fundamentals change slowly.
Analysis methods depend on new statistical developments or theory. These in turn depend on (i) computing hardware, (ii) reliable and accessible software, (iii) training and re-training of trialists in the use of new methods, (iv) acceptance of the procedure by the statistical and biological communities, and (v) sufficient time for the innovations to diffuse into practice.

It is equally important to understand what changes or new concepts do not improve methodology, but are put forward in response to nonscience issues or because of creeping regulation. A good example of this is the increasing sacrifice of expertise in favor of objectivity in clinical trial monitoring (discussed in Chapter 18). Such practices are sometimes as ill considered as they are well meaning, and may be promulgated by sponsors without peer review or national consensus.

Good trial design requires a willingness to examine many alternatives within the confines of reliably answering the basic biological question. The most common errors related to trial design are devoting insufficient resources or time to the study, rigidly using standard types of designs when better (e.g., more efficient) designs are available, or undoing the benefits of a good design with poor execution or analysis. I hope that the reader of this book will come to understand where there is much flexibility in the design and analysis of trials and where there is not.

1.4 OTHER SOURCES OF KNOWLEDGE

The periodical literature related to clinical trial methods is large and can only be approached via electronic searches. The relevant journals are not all accessible through a single source. There is also a considerable volume of methodology appearing in clinical journals. Some useful web resources are listed in Table 1.1. With regard to the Internet, it is important to recognize that it is not static, and therefore not the best source for references. However, many documents important to clinical trials can be found there with simple searches. Some of these documents in text form have been dropped from this edition because of the ease with which they can be located online.

A number of books and monographs have dealt with many facets of clinical trials. I will mention only a few that have been recently updated or have proved to be quite durable. The classic text by Meinert [1026, 1028] is recently updated and takes a practical view of the organization, infrastructure, and administrative supports necessary to perform multicenter randomized trials. In fourth edition as of 2010 is the book by Friedman, Furberg, and DeMets [546] that is an excellent introduction. Pocock [1216] also discusses conceptual and practical issues without the need for extensive statistical background, and has a recent series of design papers oriented to the clinician investigator [1220–1223]. Some recent design developments are covered by Harrington [678], and controversial topics are discussed by Chow [258]. Every trialist should read the extensive work on placebos by Shapiro and Shapiro [1367]. Monographs about clinical trials appear occasionally in disease-specific contexts. This was true in cancer and AIDS in the last decade or two, but many of the books are now slightly dated.

Even in a very active program of clinical research, a relatively short exposure to the practical side of clinical trials cannot illustrate all the important lessons. This is because it may take years for any single clinical trial, and many such studies, to yield all of their
TABLE 1.1  Some Web Resources for Clinical Trials Information

<table>
<thead>
<tr>
<th>Website</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>assert-statement.org</td>
<td>Standards for scientific and ethical review of clinical trials</td>
</tr>
<tr>
<td>cochrane.org</td>
<td>Trial-based information about the effects of health care</td>
</tr>
<tr>
<td>clinicaltrials.gov</td>
<td>Federally and privately supported clinical research</td>
</tr>
<tr>
<td>consort-statement.org</td>
<td>Evidence-based tool to improve the quality of reports of randomized trials</td>
</tr>
<tr>
<td>jameslindlibrary.org</td>
<td>Evolution of fair tests of medical treatments; examples including key passages of text</td>
</tr>
<tr>
<td>icmje.org</td>
<td>Uniform requirements for manuscripts submitted to biomedical journals</td>
</tr>
<tr>
<td>gpp-guidelines.org</td>
<td>Encourages responsible and ethical publication of clinical trials sponsored by pharmaceutical companies</td>
</tr>
<tr>
<td>mcclurenet.com/ICHefficacy.html</td>
<td>ICH efficacy guidelines</td>
</tr>
<tr>
<td>controlled-trials.com</td>
<td>Current controlled trials: provides access to peer-reviewed biomedical research</td>
</tr>
</tbody>
</table>

information useful for learning about methodology. Even so, the student of clinical trials will learn some lessons more quickly by being involved in an actual study, compared with simply studying theory. In this book, I illustrate many concepts with published trials. In this way, the reader can have the benefit of observing studies from a long-term perspective, which would otherwise be difficult to acquire.

1.5  NOTATION AND TERMINOLOGY

There is no escaping the need for mathematical formalism in the study of clinical trials. It would be unreasonable, a priori, to expect mathematics to be as useful as it is in describing nature [1555]. Nevertheless, it is, and the mathematics of probability is the particular area most helpful for clinical trials. Galileo said:

The book of the universe is written in mathematical language, without which one wanders in vain through a dark labyrinth.

Statisticians light their dark labyrinth using abstract ideas and symbols (Greek letters among them) as a shorthand for important mathematical quantities and concepts. I will also use these ideas and symbols when appropriate in this book. Mathematics is a superb shorthand, and many essential ideas cannot be explained well without its notation. However, because this is not a book primarily about statistics, the use of symbols and abstraction is as minimal as possible. A review and explanation of common usage of symbols consistent with the clinical trials literature is given in Appendix B. Because some statistical terms may be unfamiliar to some readers, definitions and examples are also listed in that appendix. Abbreviations used in the book are also explained there.

This book cannot provide the fundamental statistical background needed to understand clinical trial design and analysis thoroughly. As stated above, I assume the reader already
has much of this knowledge. Help with statistical principles that underlie clinical trial methods is available in the form of many practical and readable references for “medical statistics.” I would add only that some older texts dealing with that topic may be perfectly satisfactory since the fundamentals change quite slowly. Some specialized references will be mentioned later.

1.5.1 Clinical Trial Terminology

The terminology of clinical trials is not without its vagaries, despite Meinert’s efforts to standardize definitions in a dictionary devoted to the topic [1029]. Most of my terms are consistent with those definitions (also see Appendix B). However, I employ descriptive alternatives to the widely used, uninformativeness, inconsistent, outdated, and difficult-to-generalize “phase I, II, III, or IV” designations for clinical trials. That opaque terminology seems to have been derived or solidified from the heavy use of developmental trials for cytotoxic drugs in the closing decades of the twentieth century. Those terms have become jargon and found their way inappropriately to other contexts, helped by regulatory speak. In research and clinical care, we would never permit imprecision to the degree allowed in this old clinical trial jargon. It really can be bad enough to be unethical, and my belief is that such jargon is presently inhibiting creative design.

Drug development terms can be ambiguous or inappropriate for nondrug trials. Even newer cancer therapies may not fit easily into the old terminology. Although medical disciplines tend to view their own research issues as being unique, and encourage local or specialty-specific terminology, teaching requires terminology that is independent of the context. Some terms that I specifically avoid are pilot, proof of concept, and exploratory. These terms tend to be context dependent, but more importantly have no definitions. They seem to be used mainly to deflect criticisms from obvious design flaws.

1.5.2 Drug Development Traditionally Recognizes Four Trial Design Types

One cannot separate terminology from the developmental paradigm employed. This actually represents the entire issue at hand, because incorrect terms imply an incorrect paradigm. A look at Chapter 10 might be useful now because there the paradigm and terminology are consistent and descriptive. In therapeutic drug (especially cytotoxic drug) development historically, clinical trials were classified simply as phases I, II, III, and IV. Phase I studies are pharmacologically oriented and usually attempt to find the best dose of drug to employ. Phase II trials look for evidence of activity, efficacy, and safety at a fixed dose. Phases I and II are usually not formally hypothesis driven, meaning that comparisons to other treatments are often external to the experiment.

In phase III, new treatments are compared with alternatives, no therapy, or placebo. The comparison group is internal to the design. Investigators would not undertake the greater expense and effort of phase III comparative testing unless there was preliminary evidence from phase I and II that a new treatment was safe and active or effective. Phase IV is postmarketing surveillance and may occur after regulatory approval of a new treatment or drug to look for uncommon side effects. This type of study design is also used for purposes other than safety and activity, such as marketing, and to uncover potential new indications that might support continued product exclusivity.
Although this terminology has been widely applied to therapeutics, it often does not fit well. In some cancer settings, phase II is divided into IIa and IIb. Phase IIa trials are small-scale feasibility studies using surrogate or intermediate endpoints such as cancer precursor lesions or biomarkers. Surrogate outcomes are defined and discussed in Chapter 5. Phase IIb trials are randomized comparative studies using intermediate endpoints. Phase III cancer prevention trials are comparative designs (like IIb) using definitive clinical endpoints such as cancer incidence [831]. Some cancer prevention investigators have used the term “phase IV” to mean a defined population study [641, 642]. The same authors define phase V to be demonstration and implementation trials.

### 1.5.3 Descriptive Terminology Is Better

The “phase I, II, III” labels have become too strong a metaphor, and frequently interfere with our need to think creatively about study purposes and designs. It often seems fitting that the labels have Roman numbering. It is common to see protocols with “phase I/II” titles, and I have even seen one or two titled “phase I/II/III,” indicating how investigators labor under the restrictive labels to accommodate more flexible goals (or how confused they are).

A more general and useful description of studies should take into account the purposes of the design and the stage of the developmental paradigm, independent of the treatment being studied. A mapping from old to new terminology is provided in Table 1.2. Although mostly obvious, the descriptions are unconventional, so I will use them in parallel with traditional labels when needed for clarity. In this book, the terms phase I, II, III, and IV, if used at all, refer narrowly to drug trials.

Descriptive terminology also accounts for translational trials, an important class discussed in Chapter 11. It distinguishes several types of early developmental designs, particularly those aimed at establishing a safe dose of a new drug or agent to study. Comparative efficacy is an important class of trial designs that embody many of the rigorous design fundamentals discussed throughout this book. Details about these design types are given in Chapter 14.
1.6 EXAMPLES, DATA, AND PROGRAMS

It is not possible to learn all the important lessons about clinical trials from classroom instruction or reading, nor is it possible for every student to be involved with actual trials as part of a structured course. This problem is most correctable for topics related to the analysis of trial results, where real data can usually be provided easily. For some examples and problems used in this book, data are provided on a Web site described in Appendix A. I have made a concerted effort to provide examples of trials that are instructive but small, so as to be digestible by the student. Computerized data files and some programs to read and analyze them are provided. The site also contains some sample size and related programs that may be helpful for design calculations. More powerful sample size (and other) design software is available commercially and elsewhere on the Internet.

Experience with the design of clinical trials can be difficult to acquire. Although oriented primarily toward that topic, this book cannot replace an experienced teacher, good mentorship, collaborations with knowledgeable experts, and participation in many trials. I encourage the student to work as many examples as feasible and return to the discussion here as actual circumstances arise.

1.7 SUMMARY

The premise of this book is that well-designed experimental research is a necessary basis for therapeutic development and clinical care decisions. The purpose of this book is to address issues in the methodology of clinical trials in a format accessible to interested clinical trialists and statistical scientists. The audience is intended to be practicing clinicians, statisticians, trialists, and others with a need for understanding good clinical research methodology. The reader familiar with clinical trials will notice a few differences from usual discussions, including descriptive terms for types of trials, and an even-handed treatment of different statistical perspectives. Examples from the clinical trials literature are used, and data and computer programs for some topics are available. A review of essential notation and terminology is also provided.