Success in Academic Surgery
Series Editors: Herbert Chen · Lillian Kao

Timothy M. Pawlik
Julie A. Sosa *Editors*

Success in Academic Surgery: Clinical Trials

Springer

Success in Academic Surgery

Series Editors
Herbert Chen
Lillian Kao

For further volumes:
http://www.springer.com/series/11216

Timothy M. Pawlik • Julie A. Sosa
Editors

Success in Academic Surgery: Clinical Trials

Springer

Editors
Timothy M. Pawlik
Department of Surgery
Johns Hopkins Hospital
Baltimore, MD
USA

Julie A. Sosa
Department of Surgery
Duke University School of Medicine
Durham, NC
USA

ISBN 978-1-4471-4678-0 ISBN 978-1-4471-4679-7 (eBook)
DOI 10.1007/978-1-4471-4679-7
Springer London Heidelberg New York Dordrecht

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

We dedicate this book to our families and our colleagues, who consistently give us the inspiration to seek new knowledge. We also dedicate this book to patients everywhere who participate in clinical trials in their pursuit of hope, knowledge, and discovery.

Foreword

This book is long overdue as a reference source for surgeons conducting clinical trials and outcomes research. It has been some years since such a comprehensive and up-to-date treatise has been available for those in the arena of conducting clinical studies or those aspiring to do so.

As a profession, surgical practice has evolved in the past few decades from an empirically based profession to one with a scientific underpinning based on prospective evidence that defines an increasing proportion of our surgical practice. Yet, so many facets of surgical practice are advancing so rapidly, such as in the fields of surgical oncology and transplantation, that it is becoming increasingly difficult for the practicing surgeons to know how to incorporate these new facts and technologies into their surgical practice. This can only be accomplished with consistency across practices if we have a common factual basis for our treatment recommendations through the scientific rigor of clinical trials.

I like the way this book is organized, especially for the neophyte clinical investigator. Clinical research these days is much more complicated and regulated than ever before, so it takes a greater degree of understanding about the organizational structure, including team-building, as well as data management and ethics of patient research. Likewise, I like the emphasis on the transfer of new knowledge through a thoughtful and informative scientific publication, for I have witnessed as a journal editor, the tragedy of reviewing a manuscript with potentially important trial outcomes that had to be rejected because of poor composition.

Paradoxically, the scientific and technical advances applicable to the surgical patient have expanded dramatically, at the same time that the environment and funding for clinical research has become more arduous than ever before. Yet, the need for documenting systematically how to incorporate new devices, diagnostics, and drugs into surgical practice has never been greater. All of us in the surgical community, whether in an academic or private practice setting, must commit to participating in clinical trials as an integral component of our practice. The stakes are high and the urgency of conducting clinical trials as the scientific basis for our profession has never been greater! This valuable treatise is a practical and important manual on how to implement that commitment in our daily surgical practice.

Dallas, TX, USA — Charles M. Balch, MD, FACS

Preface

Clinical trials play a central role in the understanding of disease processes, as well as the identification of new therapeutic interventions. They help the scientific community find better ways to care for patients – whether it be through prevention, detection, or treatment of diseases. Unlike other types of scientific study, clinical trials are unique, in that the research specifically tests the given hypothesis in people. Participants in clinical trials represent a broad spectrum of individuals – healthy controls, patients with advanced stage disease, as well as patients of different ages, ethnic groups, and genders. What unifies study subjects, however, is the trust they place in the clinical investigator. As such, investigators involved in clinical trials have the professional responsibility to carry out their research not only with the highest ethical standards, but also with scientific and methodological expertise. Through ensuring rigorous clinical trial study design and implementation, investigators fulfill the contract they have with their study participants.

This book aims to equip young investigators interested in performing clinical trials with the basic fund of knowledge and tools they will need to start their careers as academic surgeons focused on clinical research. It is a primer that covers a broad spectrum of topics important to the clinical trials investigator, including the history of clinical trials, statistical considerations, regulatory issues, data management, presentation and publication of results, and ethics. While additional training and experience is needed beyond the scope of this book for those academic surgeons who plan to make clinical trials a significant part of their professional lives, the book provides important information to those seeking to know more about clinical research.

It is our hope that the book will help to cultivate the next generation of clinical scientists who will continue to seek new knowledge to improve patient care.

Baltimore, MD — Timothy M. Pawlik, MD, MPH, PhD
Durham, NC — Julie Ann Sosa, MD, MA

Acknowledgments

We would like to acknowledge the Association for Academic Surgery for supporting the *Success in Academic Surgery* series of books, as well as the efforts of all the young academic surgeons who have acted as our motivation to undertake this effort.

Contents

Contributors

H. Richard Alexander Jr., MD Division of General and Oncologic Surgery, The Department of Surgery, Marlene and Stewart Greenebaum Cancer Center, University of Maryland School of Maryland, Baltimore, MA, USA

Division of General and Oncologic Surgery, Department of Surgery, University of Maryland School of Medicine, Baltimore, MA, USA

Peter Angelos Department of Surgery, University of Chicago, Chicago, IL, USA

Christina E. Bailey, MD, MSCI Department of Surgical Oncology, University of Texas, MD Anderson Cancer Center, Houston, TX, USA

Faisal G. Bakaeen, MD Cardiothoracic Surgery, Baylor College of Medicine, Michael E DeBakey Veterans Affairs Medical Center, Houston, TX, USA

Anees B. Chagpar Department of Surgery, Yale University School of Medicine, New Haven, CT, USA

George J. Chang, MD, MS Department of Surgical Oncology, University of Texas, MD Anderson Cancer Center, Houston, TX, USA

Lorraine D. Cornwell, MD Cardiothoracic Surgery, Baylor College of Medicine, Michael E DeBakey Veterans Affairs Medical Center, Houston, TX, USA

Marquita Decker, MD, MPH Department of Surgery, University of Wisconsin, Madison, WI, USA

Shannon Decker, MPH, JD Division of Medical Oncology, The Department of Medicine, Marlene and Stewart Greenebaum Cancer Center, University of Maryland School of Maryland, Baltimore, MA, USA

Julie A. Freischlag, MD Department of Surgery, Johns Hopkins Hospital, Baltimore, MD, USA

T. Clark Gamblin, MD, MS Department of Surgery, Medical College of Wisconsin, Milwaukee, WI, USA

Jessica E. Gosnell, MD Department of General Surgery, UCSF Mt Zion Medical Center, San Francisco, CA, USA

University of California, San Francisco, CA, USA

Kamran Idrees, MD Division of Surgical Oncology and Endocrine Surgery, Department of Surgery, Vanderbilt University, Nashville, TN, USA

Lisa Jacobs, MD Department of Surgery, Johns Hopkins University, Baltimore, MA, USA

Brigid K. Killelea Department of Surgery, Yale University School of Medicine, New Haven, CT, USA

Daniela P. Ladner, MD, MPH Division of Transplantation, Department of Surgery, Northwestern University Transplant Outcomes Research Collaborative (NUTORC), Chicago, IL, USA

Department of Surgery, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA

Mahmoud Malas, MD, MHS Department of Vascular and Endovascular Surgery, Johns Hopkins Bayview Medical Center, Baltimore, MD, USA

Lisa M. McElroy, MD Division of Transplantation, Department of Surgery, Northwestern University Transplant Outcomes Research Collaborative (NUTORC), Chicago, IL, USA

Department of Surgery, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA

Nipun Merchant, MD Division of Surgical Oncology and Endocrine Surgery, Department of Surgery, Vanderbilt University, Nashville, TN, USA

Jukes P. Namm Department of Surgery, University of Chicago, Chicago, IL, USA

Shuab Omer, MD Cardiothoracic Surgery, Baylor College of Medicine, Michael E DeBakey Veterans Affairs Medical Center, Houston, TX, USA

Babak J. Orandi, MD, MSc Department of Surgery, Johns Hopkins Hospital, Baltimore, MD, USA

William S. Ragalie, MD Department of Surgery, Medical College of Wisconsin, Milwaukee, WI, USA

Sanziana Roman, MD, FACS Department of Surgery, Duke University School of Medicine, Durham, NC, USA

Edward Sausville, MD, PhD Division of Medical Oncology, The Department of Medicine, Marlene and Stewart Greenebaum Cancer Center, University of Maryland School of Maryland, Baltimore, MA, USA

Yu Shyr, PhD Department of Biostatistics, Vanderbilt Center for Quantitative Sciences, Nashville, TN, USA

Lee Gravatt Wilke, MD, FACS Department of Surgery, University of Wisconsin, Madison, WI, USA

Sandra L. Wong, MD, MS Department of Surgery, University of Michigan, Ann Arbor, MI, USA

Fei Ye, PhD Division of Cancer Biostatistics, Department of Biostatistics, Vanderbilt Center for Quantitative Sciences, Nashville, TN, USA

Abbreviations

CFR	Code of Federal Regulations
CONSORT	Consolidated Standards of Reporting Trials
DHHS	Department of Health and Human Services
DSMB	Data and Safety Monitoring Board
FDA	Food and Drug Administration
GCP	Good Clinical Practice
ICMJE	International Committee of Medical Journal Editors
IRB	Institutional Review Board
IRB	Institutional Review Board
NAACP	National Association for the Advancement of Colored People
RCT	Randomized Controlled Trial

Chapter 1
Building Your Clinical Trial Research Team

Jessica E. Gosnell

1.1 Building Your Clinical Trial Research Team

> On the 20th of May 1947, I took 12 patients in the scurvy, on board the Salisbury at sea. Their cases were as similar as I could have them. They all in general had putrid gums, the spots and lassitude, with weakness of their knees…Two of these were ordered each a quart of cyder a-day. Two otheres took two spoon-fuls of vinegar three times a-day. Two otheres took 25 gutts of elixer vitriol three times a-day. Two of the worse were put under a course of sea-water…Two otheres had each two oranges and one lemon given them every day. These they ate with great greediness…The two remaining patients took the bigness of nutmeg three times a-day.
>
> The consequence was, that the most sudden and visible good effects were perceived from the use of the oranges and lemons; one of these who had taken them, being at the end of 6 days, fit for duty.
>
> James Lind [9]

One of the earliest reported clinical trials involved the treatment of scurvy [9]. In this streamlined clinical trial, one study doctor, at sea, developed a protocol with 12 patients and 6 interventions and published findings that eventually led to the cure of this previously fatal disease. Clinical trials are of course much more complex today, with detailed protocols, highly regimented record keeping and data reporting, independently hired vendors and study monitors, high ethical standards, potential conflicts of interest, and a web of federal state and institutional regulating bodies. Recent changes in the health care environment in the United States have added to the challenges [13]. Certainly, more than one study doctor is

J.E. Gosnell, MD
Department of General Surgery, UCSF Mt Zion Medical Center,
1600 Divisadero Street, 1674, San Francisco, CA 94143, USA

University of California, San Francisco, CA, USA
e-mail: jessica.gosnell@ucsfmedctr.org

T.M. Pawlik, J.A. Sosa (eds.), *Success in Academic Surgery: Clinical Trials*,
Success in Academic Surgery, DOI 10.1007/978-1-4471-4679-7_1,

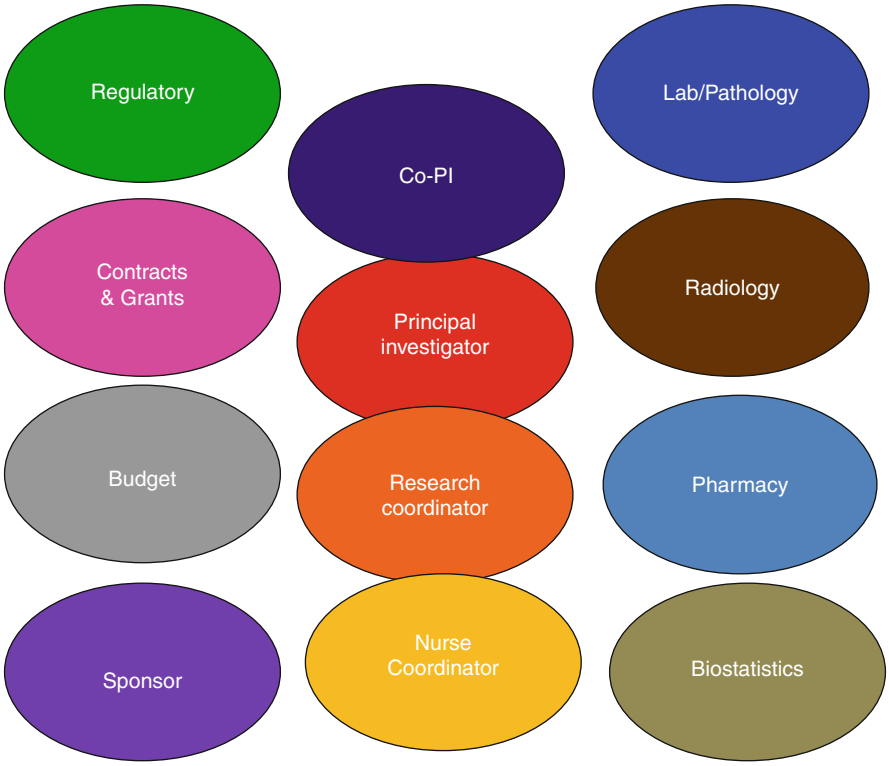


Fig. 1.1 Members of the clinical trial research team (*IRB* institutional review board, *Co-PI* coprincipal investigator)

usually needed. In fact, many people may be needed, preferably with very different but complementary skill sets. For these reasons, a dedicated clinical trial research team is essential.

Two key elements in the successful conduct of a clinical trial are assembling the clinical trial research team (Fig. 1.1) and defining the roles of each member. Developing this infrastructure early will help avoid innumerable problems during the conduct of the trial. It will also protect patients and foster good science. Below is a description of the members of the clinical trial research team and how and why each is important. There is considerable overlap between some members of the research team, and various tasks can be performed by multiple members of the research team. The research team is meant to be adaptive; it can retract to fewer members for smaller clinical trials and expand to use all listed members in larger clinical trials. The main portion of the chapter will focus on defining the various members of the clinical trial research team and their roles. I will then discuss how to build the research team, depending upon the type and scale of the clinical trial.

1.2 Key Members of the Clinical Trial Research Team

1.2.1 Principal Investigator

The principal investigator (PI) oversees, and is ultimately responsible for, all aspects of the clinical trial. As specifically outlined in the Code of Federal Regulations (21 CFR 312.60), "an investigator is responsible for ensuring that an investigation is conducted according to the signed investigator statement, the investigational plan, and applicable regulation; for protecting the rights, safety and welfare of subjects under the investigator's care; and for the control of drugs under investigation" [12]. As outlined in Table 1.1, the PI is integrally involved in all aspects of the clinical trial, from reviewing the scientific merit of the study, deciding if the trial is appropriate for the study site and the patient population, developing the budget, obtaining approval from the appropriate governing bodies (Scientific Review Committee (SRC), Investigational Review Board (IRB)), screening and enrolling patients, conducting study visits, entering data, coordinating care with the patients' clinical care team, and monitoring safety. Ideally, the PI should have a focused clinical and research interest in the clinical question being asked. He or she should have a sufficiently robust clinical practice from which to screen patients. Indeed, some have suggested that one of the most important considerations when selecting a PI is the availability of the appropriate study participants. Experience in clinical trials is obviously extremely helpful, but perhaps not essential. Well-known scientists who do not have sufficient time to devote to the trial may not be successful. As such, it is crucial that the PI assemble a clinical trial research team. While oversight is required, PIs that have to handle

Table 1.1 Responsibilities of the principal investigator

Reading and understanding the investigator brochure
Obtaining IRB approval and complying with all decisions
Conducting the study in accordance with the protocol
Maintaining adequate records and making them available for review
Personally conducting the study or supervising other trained personnel
Ensure that research team members are qualified and appropriately trained and retrained
Informing study participants about the investigational nature of clinical trials
Ensuring that informed consent is properly obtained
Reporting adverse events in accordance with federal and institutional code
Reporting any change in the research
Reporting protocol violations
Making no change in the protocol without the prior approval of the IRB

all of the tasks and procedures in the clinical trial may not have the appropriate time and perspective for true oversight and may not be able to enroll a sufficient number of patients to support the trial. Teamwork and delegation are essential. This is especially important for clinical trials that have complex protocols and are data intensive. Interestingly, in a study published by Shea and colleagues, factors associated with enrollment of eligible participants with documented myocardial infarction were reported. Possible correlations were found in institutions in which study participants were cared for by staff other than attending physicians and with the presence of a committed nurse coordinator [15].

1.2.2 Co-principal Investigators

It may be important to have additional co-PIs involved in the oversight of the clinical trial. They may be clinical partners that can help to identify potential study subjects or be a part of a larger multidisciplinary group. They may have expertise in slightly different clinical or research areas. Often, expertise in other specialities may be needed for consultation regarding risk. For example, if a study drug is known to have cardiac side effects, IRBs may require that a cardiologist review the protocol and be a formal member of the research team. Another collaborative co-PI to consider in the study team may be a radiologist or pathologist, brought on to read x-rays or review cytopathology for a clinical trial. This approach leads to less variability in the interpretation of the data but a greater potential for bias.

1.2.3 Clinical Research Coordinator/Data Manager

The clinical research coordinator is in many ways the engine of the clinical trial. The importance of their role cannot be overemphasized. The research coordinator helps prepare the informed consent and the clinical trial submission to local IRBs and interacts with patients during the screening, enrollment, and study visits. They enter data and organize study documents. They interface with industry, contracts and grants, and study monitors. They submit invoices for billing and facilitate communication with primary care physicians (www.clinicaltrialguru.com). An excellent clinical research coordinator is a true prize. Some of the more critical traits in a research coordinator include the following: organizational skills, attention to detail, interpersonal skills, and computer and database experience. A medical or scientific background may be helpful, but it is not essential and may be easier taught than the other more innate personal characteristics. It is desirable that the job be held by someone who can commit to a reasonable length of time (e.g., 1–2 years, such that multiple RCs are not needed throughout the life of the trial). Suffice to say, it is worthwhile spending considerable time and effort in hiring a research coordinator. They can make or break the clinical trial.

1.2.4 Nurse Coordinator/Research Nurse

There is some overlap between a clinical research coordinator and a nurse coordinator. In a complex clinical trial, both roles may be essential to complete the procedures, regulatory requirements, and data entry outlined in the protocol. As the clinical research coordinator focuses on the regulatory aspects of the trial, the nurse coordinator can focus on the clinical "arm" of the study. He or she can assist in screening and enrollment of study participants, both of which are data intensive. Nurse coordinators are extremely important in clinical trials involving multiple study procedures, such as electrocardiograms, venipunctures, and hemodynamic testing. Depending on their level of training, research coordinators may not be able to perform these tests.

At some institutions with well-developed infrastructure for clinical research, either or both the clinical research coordinator and the nurse coordinator may be centralized positions, in which they are involved in multiple studies that do not require 100 % time.

1.2.5 Biostatistics

A biostatistician should be an integral and early member of the clinical trial research team [11]. He or she must review the design and methodology of the clinical trial and ensure that the study is adequately powered to answer the proposed question [14]. They perform statistical analysis when the data have been collected and may be involved in interim analysis of the data. Clearly, this is crucial to the eventual success of the trial. Many academic institutions have skilled biostatisticians that can be consulted.

1.2.6 Financial

A budget is essential to ensure that a clinical trial can be conducted effectively and safely and in the black. Therefore, a financial expert needs to be identified to carefully prepare and review the budget. In-depth discussion of preparing the budget is outside the scope of this chapter. However, in general, the investigators and the sponsor must take into account (1) what the cost for study procedures are, (2) what procedures will be covered by a third-party payer, and (3) what constitutes direct and indirect costs in the study [1]. The financial expert also has the responsibility of documenting and justifying expenditures for external or extramural funding sources. In an investigator-initiated clinical trial, the financial expert may be the PI of the study. However, if the PI does not have sufficient experience in preparing budgets, it may be a senior partner or another colleague with this particular skill set. In

industry-sponsored trials, the budget may be proposed by the sponsor and then reviewed and adjusted by the PI. It is important for the budget to be as comprehensive as possible, with frequent and careful review of the protocol. Depending on the complexity of the trial and ongoing requirements of time and energy, it may be appropriate to budget in a percentage of the PI salary or support for a nurse or data coordinator.

1.2.7 Contracts and Grants

A contract is a legal agreement between the study investigator and the funding source (industry, grant organization). It incorporates the elements of the budget and helps protect both parties. It includes the management and control of data and defines publication rights and intellectual property [5, 7]. Someone intimately familiar with industry contracts, medical center or university policies, and the specific protocol under study should be contacted to review this document. This person may be identified (or assigned to your clinical trial) through a number of various offices, such as Contracts and Grants Office, Regulatory Affairs Office, or Protocol or Research Administration Office [8]. A lawyer may or not be needed. The importance of this member of the clinical trial research team cannot be overemphasized, as it is difficult to renegotiate a poorly prepared contract.

1.3 Clinical Affiliates

1.3.1 Pharmacy

As many clinical trials involve interventions with medications, a pharmacist is an important part of the clinical trial research team. The pharmacist is responsible for storing, tracking, and dispensing the study drug. They must adhere to the Code of Federal Regulations, which strictly mandates the policies and procedures around study drugs. They may also be responsible for the manufacture of a placebo in placebo-controlled clinical trials.

1.3.2 Labs/Imaging/Pathology/Electrocardiogram

Many clinical trials involve obtaining and interpreting laboratory studies, radiographic imaging, and/or pathologic specimens. It is important to identify point people within these departments, both to help refine the protocol in the designing stage and for ongoing input during the clinical trial. This is particularly important for studies involving the imaging or pathology, in which interpretation can vary between practitioners. For

multicenter and/or industry-sponsored clinical trials, a number of study procedures, such as labs, x-ray, and pathology, are contracted out to independent vendors. These vendors are theoretically unbiased and independent and involve fewer individuals reading the films and slides. If multiple study vendors are being used, it is important to carefully review the study protocol, especially with regard to the processing of laboratory specimens, the preparation of imaging CDs for shipping, and the transmission of electrocardiograms. This can be a very labor-intensive component of the study and should be factored into decisions about personnel, work flow, and budgeting.

1.4 Regulatory

1.4.1 Scientific Review Board

Many institutions require that clinical trial proposals are evaluated for scientific merit prior to review by the institutional review board. These review boards must have at least three members. They have various names, including Cancer Center Scientific Review Committee and Protocol Review Committee. There may be separate committees for stem cell research, recombinant DNA, and others.

1.4.2 Institutional Review Board

All clinical trials involving human subjects must be approved by an institutional review board (IRB). IRBs review all aspects of the study, usually in a prepared paper or online application. The application includes details of the protocol, study procedures, key personnel, biostatistics, assessment of the risks and benefits of the study intervention, subject confidentiality, any financial conflicts of interest, and the informed consent document [4]. A study cannot begin until IRB approval has been obtained. Initial review may prompt questions about the clinical trial, and these must be clarified prior to approval. Of note, many institutions will have assigned people within the IRB that work with specific departments or review certain types of clinical trials. Identifying this person during the submission process can help immeasurably, as they can often anticipate potential problems that can be clarified in the first submission.

1.4.3 Data and Safety Monitoring Board

For large randomized clinical trials involving more than minimal risk to study participants, a Data and Safety Monitoring Board (DSMB) is often required. A DSMB is a group, independent from both the sponsor and the study investigators,

who monitors the clinical trial over time [6]. At least one member is usually a statistician. The board will convene at predetermined intervals and review two main components: (1) adverse events and (2) interval interpretation of the data. If there are unanticipated adverse events or more than expected, the data monitoring committee may suggest changes in the protocol or a revised consent process. Interval interpretation of the data is important in comparative or placebo-controlled studies, in which one intervention shows advantage early in the clinical trial. The DSMB has the power to recommend termination of the study based upon their findings of safety concern, futility, or outstanding benefit.

1.4.4 The Office of Human Research Protections

The Office of Human Research Protection is under the National Institute of Health and essentially oversees all the federally funded IRBs in the United States performing research on human subjects. The chair(s) of the IRB signs a Federal-Wide Assurance document, which formalizes the institution's agreement to conduct its human research in accord with federal statutes [4].

1.4.5 Food and Drug Administration

The Food and Drug Administration (FDA) is an agency of the US Department of Health and Human Services and is responsible for protecting and promoting public health through the regulation of food safety, tobacco, and pharmaceutical drugs, among others (www.fda.gov) [17]. One of their major jobs is the regulation of the safety and efficacy of investigational new drugs and devices. New drugs, or already approved drugs used for a different indication, require an Investigational New Drug (IND) application be filed with the FDA. The FDA also audits IRBs and study sites periodically.

1.5 Sponsors

1.5.1 Intra- and Extramural Grants

Funding for clinical trials can be characterized as intramural when it is within the bounds of an institution. Extramural grants are outside the institution and include private foundations, nonprofit organizations, federal organizations, and industry.

1.5.2 The National Cancer Institute

The National Cancer Institute (NCI) was established in 1937 and is the "US government's principal agency for cancer research" (www.cancer.gov) [10]. It is part of the National Institute of Health, 1 of 11 agencies that make up the US Department of Health and Human Services. The NCI supports research and training in the diagnosis, prevention, and treatment of cancer and includes both intramural and extramural funding. Infrastructure support is also available on Cancer Center Support Grants (www.cancer.gov). Finally, NCI keeps a list of ongoing clinical trials and whether they are still enrolling patients.

1.5.3 Industry

Approximately 70 % of clinical research in the United States is now funded by the private sector [2, 8]. Other potential funding sources include institutional grants, federally funding through the National Institute of Health. Clearly, the sponsor of the study will be an important part of the clinical trial research team. They develop the protocol and have control of the data. They may have protocol changes and amendments, all of which have to be submitted to the IRB for approval. Some sponsors may have contractual restrictions on publications and authorship.

1.6 Building the Clinical Trial Research Team

Clinical trials range from relatively straightforward studies, like that performed by James Lind in 1573, to cooperative, large-scale group trials like those funded by the NCI. Because the workforce needed to complete these studies is so different, the research team also needs to be adaptive—to retract or expand to support the scope of the trail. At a minimum, a clinical trial needs a fully invested principal investigator, and the clinical trial needs to be approved by the governing bodies [3, 16]. As clinical trials increase in their complexity, additional team members such as the research coordinator and/or the nurse coordinator are usually also needed to efficiently perform the regulatory submissions, study procedures, and data entry [3]. Support for the budget, contracts, and statistics is essential. Point people in the laboratory, imaging, and pathology are added as indicated by the study protocol. An increasing number of clinical trials will receive funding either from the government or from industry.

References

1. Barnes M, Korn J. Medicare reimbursement for clinical trial services: understanding Medicare coverage in establishing a clinical trial budget. J Health Law. 2005;38(4):609–31.
2. Bodenheimer T. Uneasy alliance-clinical investigators and the pharmaceutical industry. N Engl J Med. 2000;342:1539.
3. Carbone PP, Tormey DC. Organizing multicenter trials: lessons from the cooperative oncology groups. Prev Med. 1991;20(1):162–9.
4. Cohen JM. Training institutional review board members. In: Institutional review and board management and function. 2nd ed. Amdur RJ, Bankert EA, editors. Sudbury: Jones and Bartlett; 2002. p. 313.
5. Finn PB. The negotiation and development of a clinical trial agreement. J Biolaw Bus. 2006; 9(2):21–7.
6. Friedman L, DeMets D. The data monitoring committee: how it operates and why. IRB. 1981;3(4):6–8.
7. Kassirer JP, Angell M. On authorship and acknowledgments. N Engl J Med. 1991; 325:1510–2.
8. Leong SP, editor. Cancer clinical trials: proactive strategies. New York: Springer; 2007.
9. Lind J. A treatise on the scurvy. London: A. Millar; 1753.
10. National Cancer Institute at the National Institute of Health [Internet]. Bethesda. Available from http://www.cancer.gov.
11. O'Neill RT. FDAs critical path initiative: a perspective on contribution of biostatistics. Biom J. 2006;48(4):559–64.
12. Ott MB, Yingling GL. Guide to good clinical practice. Tab. 1999;100:7.
13. Rusch VW. Surgeons: a future role in clinical trials? Oncologist. 1997;2(3):5–6.
14. Shafer H. Adaptive designs from the viewpoint of an academic biostatistician. Biom J. 2006;48(4):586–90; discussion 613–22.
15. Shea S, Bigger JT, Campion J, et al. Enrollment in clinical trials: institutional factors affecting enrollment in the Cardiac Arrhythmia Suppression Trial (CAST). Control Clin Trials. 1992;13:466–86.
16. Tognoni G, Alli C, Avanzini F, et al. Randomized clinical trials in general practice: lessons learned from a failure. BMJ. 1991;303:969–71.
17. US Food and Drug Administration [Internet]. Silver Spring MD. Available from http://www.fda.gov. Updated 29 Mar 2013.

Chapter 2
Clinical Trials: Ensuring Quality and Standardization

Marquita Decker and Lee Gravatt Wilke

2.1 Introduction

Clinical equipoise regarding the best surgical or medical options for a particular illness is widespread in this innovative century and helps drive the hypotheses for clinical trials and in doing so advances in evidence-based medicine. Adherence to regulations and standardized processes within a clinical trial, the focus of this chapter, offers the best means to provide a balance between scientific progress and patient safety. With uncontrolled or nonstandardized clinical trial design and implementation, it is difficult to determine if a new treatment, surgical procedure, device, or drug has made a difference in a human disease or put those who volunteered for the clinical trial at risk for complications and progressive illness. Trials done incorrectly can lead to wasted resources and lost patient confidence in the medical system. It is therefore imperative that the surgeon scientist interested in designing clinical trials understand the historical basis for the current international and national laws and regulations surrounding clinical research as well as the steps to imbed the proper standard quality components into a clinical trial.

The design, conduct, and reporting of *surgical clinical trials* in particular have been the focus of significant criticism over the past 20 years. In 1996 a simple commentary published in Lancet launched an important challenge for surgeons. In that article Richard Horton from London wrote:

"In addition to safety and efficacy studies, more pragmatic trials are needed to determine acceptability, effectiveness and efficiency by comparing new interventions with currently preferred treatments. In 1923, the medical statistician, Major

M. Decker, MD, MPH • L.G. Wilke, MD, FACS (✉)
Department of Surgery,
University of Wisconsin,
Madison, WI 53792, USA
e-mail: wilke@surgery.wisc.edu

T.M. Pawlik, J.A. Sosa (eds.), *Success in Academic Surgery: Clinical Trials*,
Success in Academic Surgery, DOI 10.1007/978-1-4471-4679-7_2,

Greenwood wrote that 'I should like to shame {surgeons} out of the comic opera performances which they suppose are statistics of operations'. Only when the quality of publications in the surgical literature has improved will surgeons reasonably be able to rebut the charge that as much as half of the research they undertake is misconceived." [1]

How then can surgeons respond to these challenges and recommendations? How can they assure that they are designing and performing high-quality, standardized research? A recent survey of surgical trialists seeks to provide insight into how to build a stronger education platform and resources for clinical trials in surgery. Jarman and colleagues from Methodist Hospital in Texas performed semi-structured interviews of 15 individuals, primarily surgeons, involved in conducting clinical research over the past decade [2]. One of their most important discoveries, which is immensely pertinent to this chapter, was the lack of knowledge surrounding the availability of "regulatory and normative" documents to guide the development and publication of a clinical trial. Between 1949 when the Nuremberg Code was released and 2008, greater than 1,000 documents have been written and/or updated regarding the ethical and methodological standards for clinical trials, the great majority of which are available on the Internet [3]. In addition to seeking out these readings, the four primary recommendations for aspiring surgical trialists from those experts interviewed were:

1. Obtain a formal education in the methods and issues of clinical trials.
2. Identify mentors.
3. Establish a network of collaborators.
4. Search for opportunities to be involved in all aspects of clinical trials [2].

Each of these recommendations is crucial towards becoming a respected surgical trialist in a clinical research arena which is increasingly complex due to the influx of innovative device technology as well as genomic information. There are multiple formal education programs from the Masters of Public Health or Masters of Clinical Research to smaller programs offered by subspecialty societies as well as the week-long clinical trials course offered by the American College of Surgeons. Two recommended textbooks for the surgical clinical trialist which delineate the steps from hypothesis generation to dissemination of results include Lawrence Friedman's *Fundamentals of Clinical Trials* (4th ed, Springer) and Stephen Hulley's *Designing Clinical Research* (3rd ed, Lippincott Williams and Wilkins). For the junior clinical investigator, mentors and collaborators should ideally include, in addition to their senior attendings, biostatisticians with experience in clinical trial design. Finally, opportunities to be involved in clinical trials should be identified during residency or as a junior attending. Even though the resident or young attending surgeon may not play a pivotal role in the design of a clinical trial, attendance at a national clinical trial meeting and in the discussions will help guide the individual's understanding of the standardized steps involved in bringing a quality clinical trial to completion.

2.2 History

Clinical trials involving human subjects have not always been performed ethically and with respect for the subjects involved. Likely the most well-known medical trial which lacked beneficence was that performed by the US Public Health Service in 1932 termed the Tuskegee Syphilis Experiment. Though not a randomized trial, the participants were misled into believing that their syphilis was being treated versus just monitored to obtain a natural history of the disease. It was not until the 1970s when the media exposed the lack of interventions that the research was terminated [4]. It was in 1947 however that the Nuremberg Code was developed because of the atrocious crimes which were committed on individuals of Jewish descent and those with mental disorders during World War II. This Code has not been adopted into law by any nation or medical association but is recognized as the foundation on which all national and international regulations for human subjects research are based [5]. The ten research principles set forth by the Nuremberg judges highlight the importance of voluntary informed consent as well as the novel right of a subject to withdraw from the research project at any time. Interestingly, 50 years prior, Sir William Osler had recommended that informed consent be included in all medical experiments [6]

Over the next half a century, additional clinical research regulations were developed and integrated into trial design to assure the protection and safety of human subjects. These include the Helsinki Declaration, which was last updated by the World Medical Association in Korea in 2008 [7], as well as the Belmont Report set forth in the late 1970s by the US Department of Health and Human Services (DHHS). This latter report defines the three ethical principles for human subjects research which are embedded into most physicians' institutional requirements for performing human clinical trials research:

1. Respect for persons: Informed consent should be obtained voluntarily from all participants.
2. A scientifically researched question is being evaluated and that the risks are acceptable in relation to the benefits.
3. Justice: Requires equitable access to the benefits of the research. Researchers must adequately offer the trials to those who may be underrepresented such as children or minorities or justify why they cannot be included [8].

The ethics of clinical trials and human subjects research is the subject of a myriad of books and publications and is unlikely to be fully covered in a single chapter. The importance of understanding the regulatory requirements for clinical trials research however cannot be overemphasized. One of the most important guides for the conduct of clinical research is the institutional review board (IRB). The mission of an IRB is to assure ethically conducted research with minimal risks, equitable participant selection, and adequate confidentiality. Members of an IRB include researchers as well as community members and lawyers. The IRB system has up until the past few years been decentralized but is undergoing revision and now

includes several "centralized" IRBs for multi-institutional trials [9]. As the central IRBs become commonly used, it is expected that the individual institutional IRBs will be able to more efficiently review and approve research projects. In addition to familiarizing oneself with the requirements for the ethical conduct of clinical research through the aforementioned readings and interaction with an IRB, a surgical clinical trialist also needs to be familiar with several standard methodological practices such as GCP, SOPs, CRFs, and AEs as well as the CONSORT statement for reporting of clinical trials, as described in the next sections of this chapter.

2.3 Good Clinical Practice

> Good Clinical Practice (GCP) is an international ethical and scientific quality standard for designing, conducting, recording and reporting trials that involve the participation of human subjects. [10]

The International Conference on Harmonization of Technical Requirements of Pharmaceuticals for Human Use (IHC) developed these standards as a quality assurance system to cover all stages of clinical trials. With growth of clinical trials internationally and globalization of medical product sales, GCP provides a means to assure sound research practices and standardized data collection allowing for generalizable measurements of clinical effectiveness of a drug or device. Compliance with GCP can also assure the public that the rights, safety, and well-being of clinical trial subjects are protected, in accordance with the Declaration of Helsinki [7]. In addition to the definitions for GCP by IHC, each of the US government branches within Health and Human Services (HHS) provides their interpretation and description of GCP. For example, the FDA "houses" their description of GCP in the Code of Federal Regulations (CFR) Title 21, while the NIH describes their GCP principles in Title 45 of the CFR [11, 12].

GCP standards require that any proposed clinical trial be reviewed and approved by an institutional review board (IRB) or independent ethics committee (IEC) prior to initiation of the study. Although the IRB and IEC are not necessarily responsible for scientific review, they are charged with the tasks of determining if a study is feasible and assuring that it will not place study subjects at undue risk. For this to be the case, there must be *equipoise* between the study intervention and the control. There must be uncertainty about whether the control or the investigational product is more efficacious. The equipoise between study arms should be described in the trial protocol, the investigator brochure, and informed consent forms. It is critical to specify the primary outcome in the protocol to describe how equipoise will be tested and efficacy will be measured. The IRB or IEC can thus determine from the description of equipoise whether or not the trial design is feasible and safe.

The IRB or IEC is not only required to review the trial protocol but also the investigator's current curriculum vitae and qualifications for conducting the trial, available safety information about the investigational product (e.g., drug, medical

device, or surgical technique), investigator brochure if one exists for the drug or device, subject recruitment procedures, informed consent form(s), written information to be provided to subjects, and information about payments and compensation available to subjects [10]. As the principal investigator preparing a clinical trial protocol for IRB review, it will be important to keep in mind the above-listed requirements, utilize institutional templates for required forms, and learn the required or previously utilized *standard operating procedures* (SOPs) for study activities pertinent to a trial. SOPs are detailed, written instructions to achieve uniformity in the performance of specific study functions [10] that should be included in the study protocol. In the conduct of a clinical trial, SOPs provide uniformity in the administration of the investigational product or implementation of the intervention. Detailed SOPs are particularly necessary in surgical clinical trials, where procedural interventions need to be uniform. For example, the Swedish Rectal Cancer Trial was criticized for not having a standardized method of surgical resection in studying the benefits of preoperative radiation therapy [13]. In contrast, the Dutch TME trial [14], which defined a standardized surgical procedure, demonstrated a clear benefit to preoperative radiation therapy after total mesorectal excision of rectal cancer and was widely accepted. This single example demonstrates that if standardized operating procedures, especially in regard to operative technique, are utilized in the conduct of a clinical trial, the results are more likely to be interpretable and generalizable.

2.4 Case Report Forms

The case report form (CRF) is the most basic unit of data collection in a clinical trial. It is a printed or electronic document (remote data capture or data entry (RDC/RDE)) designed to record all of the protocol required information to be reported to the sponsor (national, international, or industry) on each trial subject [10].When designing a clinical trial, the principal investigators need to assure that they include all data that needs to be collected in the protocol, as information that has not been specified as analytic in the protocol cannot be included on a CRF. Trialists should be thoughtful about keeping the data acquisition simple and standardized with minimal free text.

There are currently no universally standard CRF designs; however, there are several ongoing efforts to harmonize the information acquired within and across clinical trials and avoid duplication of efforts in design. For example, a subproject of the Clinical Data Interchange Standards Consortium (CDISC, http://www.cdisc.org), the Clinical Data Acquisition Standards Harmonization (CDASH), is working on standard forms for safety data as well as disease or therapeutic specific forms and data elements [15]. Some of these forms and processes for data acquisition will become outdated quickly due to the burgeoning amount of calculations and information that can be obtained from patients and human tissues. The associated field of clinical research informatics is one that is growing exponentially and is working to

provide leadership and direction on these issues of standardization of common data elements and medical terminology as well as the best methods for compilation and storage of metadata from clinical trials [15]. The National Cancer Institute (NCI) has one such example of a CRF Harmonization and Standardization Initiative to design standardized case report forms incorporating common data elements to "improve information sharing among cancer researchers and optimize data requirements in collaboration with the Federal Drug Administration (FDA)" [16]. By reducing the time spent in developing a data collection strategy for each cancer clinical trial, a core library will allow for rapid initiation of a novel clinical trial and dissemination of information across linked clinical studies (Fig. 2.1).

2.5 Adverse Event Reporting

While the CRF records all events in a trial related to each subject, a subset of the record must include adverse events that occur during and after the course of a clinical trial. *Adverse events (AE)* are defined as any unfavorable or unintended medical occurrence, including any abnormal sign (physical exam or laboratory finding), symptom, or disease temporarily associated with the use of a medical treatment whether or not it is considered related to the medical treatment (causality) [10, 12]. The purpose of AE reporting is to monitor the safety of the study, inform the investigators, regulators, and subjects of risks potentially associated with the intervention and allow for timely intervention if it appears that the risks of the clinical trial have changed. In general the majority of AE reporting has centered around events associated with drugs and devices, and less standardization has occurred with reporting of adverse events in purely surgical clinical trials. In the United States, adverse event reporting has recently been restructured to address causality more specifically [12]. With the globalization and massive growth of clinical trials over the past 10 years [17], the FDA has been flooded with volumes of adverse event reports and other regulatory documentation. Many of these adverse event reports have been uninterpretable or irrelevant as reporting individuals were often blinded and unable to relate events to the investigational drugs or devices [18]. The Clinical Trials Transformation Initiative has since worked to improve the quality and efficiency of reporting by requiring the investigational sponsor to assign attribution or causality to each adverse event and provide both the FDA and the study investigators with an aggregate report of adverse events that are attributable to the investigational drug or device.

The process surrounding the ascertainment, recording, and reporting of adverse events remains moderately challenging especially to the surgeon scientist who may not be involved in drug- or device-related surgical research. It is however important to understand the basic definitions of AE reporting in order to ensure that the proper information is included in a clinical trial protocol, so AE determination does not become challenging because adequate definitions were not provided a priori. *Expected* adverse events are those that have been identified in nature, severity, or

HEADER: PI NAME, Protocol or IRB Number, Protocol Short Title

Subject Initials ☐☐☐ **Subject ID** ☐☐☐☐☐ **Date:** ☐☐ / ☐☐ / ☐☐
Month Day Year

Demographics

Subject UWHC Medical Record Number*: ☐☐☐☐☐☐☐

First Name*:
Middle Name (or initial):
Last Name*:

Birthdate*: ☐☐ / ☐☐ / ☐☐☐☐ ☐ Approximate ☐ Unavailable
Month Day Year

Gender*: (check one)
☐ Male
☐ Female
☐ Unknown or Not Reported

Ethnicity*: (check one)
☐ Hispanic
☐ Non-Hispanic
☐ Unknown or Not Reported

Race*: (check all that apply)
☐ American Indian or Alaska Native
☐ Asian
☐ Black or African American
☐ Native Hawaiian or Other Pacific Islander
☐ White or Caucasian
☐ Unknown or Not Reported

Other Medical Record Number(s):

Medical Record Number	Hospital/Care Provider (e.g. UW Hospital, VA, etc.)

Contact Information:

Address:		Unit #:
City:	State:	Zip:
Phone Number: ☐ Home ☐ Work ☐ Cell ☐ Other	Alternate Phone Number: ☐ Home ☐ Work ☐ Cell ☐ Other	Email address:
Preferred method of contact:		

Emergency Contact:

Name:		
Address:		Unit #:
City:	State:	Zip:
Phone Number: ☐ Home ☐ Work ☐ Cell ☐ Other	Alternate Phone Number: ☐ Home ☐ Work ☐ Cell ☐ Other	Email address:
Preferred method of contact:		

*indicates required field

Form Completed By: _____ **Date:** _____

Fig. 2.1 Example of case report form template

frequency in the current clinical trial protocol, investigator brochure (if available for a drug or device), and current consent form. For example, a 5 % incidence of surgical site infection is anticipated or expected in a clinical trial that tests the efficacy of laparoscopic versus open distal pancreatectomy for resection of benign pancreatic tumors. The research team is informed of this in the clinical trial protocol, and

subjects are made aware of this risk during the consent process. Therefore an occurrence of surgical site infection in a subject would be an expected event but would still be recorded on the AE forms to determine the extent of this event across the entire patient accrual. Anticipated or expected events can be internally documented, monitored, and analyzed in aggregate.

An *unexpected* adverse event is any event, the severity or specificity of which is not consistent with the risk information described in the clinical protocol or investigator brochure, and is more likely than not related to the research, and it suggests that the research places subjects or others at greater risk than previously known [19]. "Relatedness" of the AE to the clinical study is defined differently for each regulatory body:

1. ICH definition requires evidence suggesting a causal association.
2. FDA definition requires a reasonable possibility that the AE is associated with the research drug or device.
3. NCI has a more complex reporting structure for AEs which is best described below in the description of CTCAE but moves through five categories from not related to definitely related to the intervention.

It is worth noting that not all unanticipated problems involve an unexpected adverse event. For example, the absence of laparoscopic equipment for a patient randomized to the laparoscopic approach in the aforementioned distal pancreatectomy trial would be an unanticipated problem but not an unexpected adverse event as would the administration of an incorrect dose of a study medication if drawn up incorrectly by the pharmacy. The distinction between unanticipated and unexpected outcomes in the ascertainment of adverse events is important because they each prompt different documentation and management processes. Unanticipated adverse events such as incorrect study drug dosing are reported to the IRB, the patient, and potentially the sponsor but are recognized to not be causal in adverse events. Unexpected adverse events must be reported to the IRB of the study site and subsequently sent to the Office for Human Research Protection in accordance with US Health and Human Services regulations 45 CFR 46.103(a) and (b)(5) [11, 19].

Prior to management of an unexpected adverse event, some *analysis* is required to understand the process that led to the event and how to address it. First a basic classification of the adverse event is needed. The Medical Dictionary for Regulatory Activities [20] is a standardized system of medical terminology that is used internationally to classify adverse event information associated with the use of biopharmaceuticals and other medical products. MedDRA is maintained by the International Conference on Harmonization and is the required terminology in reports to many regulatory agencies. The National Cancer Institutes' Common Terminology Criteria for Adverse Events (CTCAE) [21] is also an instrument used to classify adverse events. It uses clinical and laboratory evaluation criteria and is required in studies that are funded by the NCI. Grading of adverse event severity is a key component of the classification process. The CTCAE has a grading scale of 0–5. Zero is no adverse event, 1 is mild, 2 is moderate, 3 is severe, 4 is life-threatening, and 5 is

death [21]. Life-threatening adverse events and death must be reported to the IRB and the FDA within 7 calendar days, in accordance with 21 CFR 312.32(a) [12]. Of note, the CTCAE grading scale describes *severity* of an event but not "seriousness." SAEs or serious adverse events are any adverse drug event (experience) occurring at any dose that results in ANY of the following outcomes: "death, a life-threatening adverse drug experience, inpatient hospitalization or prolongation of existing hospitalization (for >24 h), a persistent or significant incapacity or substantial disruption of the ability to conduct normal life functions, a congenital anomaly/birth defect. Important Medical Events (IME) that may not result in death, be life threatening, or require hospitalization may be considered a serious adverse drug experience when, based upon medical judgment, they may jeopardize the patient or subject and may require medical or surgical intervention to prevent one of the outcomes listed in this definition" [12]. The process of reporting SAEs is variable for clinical trials and can be anywhere from within 24 h to up to 10 days depending on the sponsor of the trial and associated regulatory agencies. For example, a phase I trial evaluating the safety of a novel drug may require the submission of an SAE within 24 h to assure that additional patients are not placed on the study until the SAE has been fully investigated.

The standard operating procedures of a clinical trial should also delineate the methods for *ascertaining* adverse events. Adverse events can be ascertained by elicited or volunteered report from subjects or by direct observation. Direct observation is the most reliable method to ascertain adverse events, but it is not always feasible. Studies have demonstrated that adverse event reporting increases when subjects are asked open-ended questions to elicit reports of adverse events [22–24]. Subject-elicited adverse events or patient-reported outcomes (PRO) are increasingly studied, particularly with the inception of patient-centered outcomes research [25, 26]. Research focused on the development of standardized and generalizable patient-reported outcome instruments remains ongoing. The majority of clinical trials currently rely on volunteered reporting of adverse events or direct observation of them during study visit examinations.

The ICH has developed the Harmonized Tripartite Guidelines [10] that detail the key components of an adverse event report, summarized here:

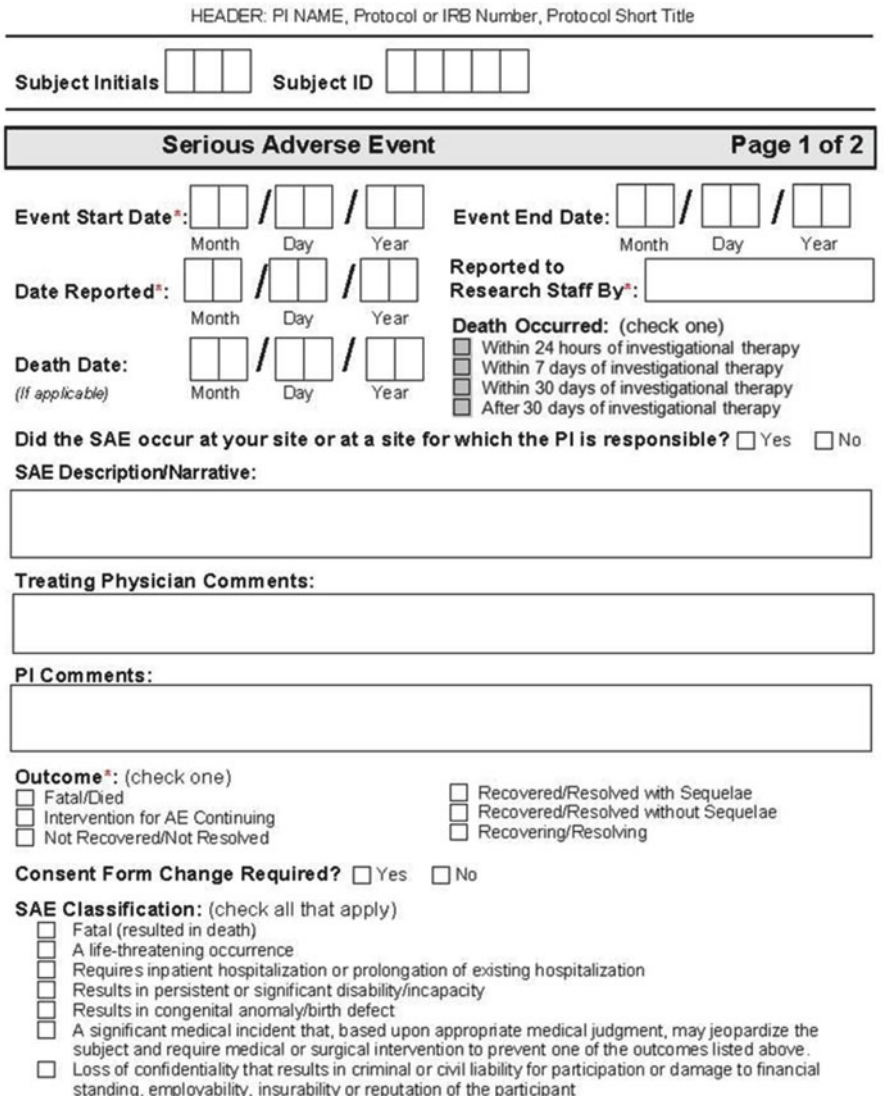
1. Subject identifier
2. Demographic data: age, race, sex, weight, and height
3. Location of case report forms and source documents (for reference)
4. The adverse event
5. Duration of the adverse event
6. Severity (mild, moderate, severe)
7. Seriousness (life-threatening, death)
8. Action taken (none, dose reduced, treatment stopped, specific treatment instituted, etc.)
9. Outcome
10. Causality assessment (unrelated, unlikely, possible, probable, definite causality)

11. Date of onset or date of clinic visit at which the event was discovered
12. Timing and onset of the adverse event in relation to last dose of test drug/investigational product
13. Study treatment at time of event or most recent study treatment taken
14. Test drug/investigational product dose at time of event
15. Drug concentration
16. Duration of test drug/investigational product treatment
17. Concomitant treatment during study

The ICH guidelines state that all adverse events for each patient, including the same event on several occasions, should be documented. Study sponsors and investigators are held responsible for the reporting of adverse events. Separate reporting guidelines exist for trials involving recombinant DNA molecules or gene transfer [27]. In these trials the principal investigator is responsible for ensuring that safety reporting requirements are fulfilled, although such reporting may be delegated to the investigational sponsor.

A recent publication by Virzi et al. from the Pediatric Heart Network (PHN) highlights the complexities of AE reporting in surgical trials [28]. In 2005, the PHN opened a randomized surgical clinical trial evaluating single ventricle reconstruction in children with hypoplastic left heart syndrome. The initial AE reporting system is commonly used for drug trials, but in this critically ill neonatal population, every lab value and event would require a serious AE report. With approval the team was able to switch to a "sentinel" event reporting system for serious AEs thereby reducing the administrative burden and highlighting the critical events of interest in this study [28].

In a qualitative review of good clinical practice (GCP) inspection citations [29], investigators and sponsors were frequently reprimanded for inconsistent adverse event documentation between case report forms, adverse event logs, and other source documents. An important lesson to take from this review is that it is best to either document adverse events consistently and concordantly in designated forms or document adverse events in one central source record. In studies that require adverse events to be reported in more than one document, it is best to reference all source documentation in each report. The review of GCP inspection citations also highlighted a frequent problem with undocumented telephone correspondence between investigators and study sponsors related to the reporting and management of adverse events. Management of adverse events must be documented as accurately and consistently as the events themselves. The best advice for young investigators interested in designing clinical trials is to link with a "seasoned" clinical research specialist who is knowledgeable on AE reporting and has had extensive experience in this process so that the clinical trial protocol can be written and the required forms included to ensure that the AE process goes smoothly and ensures that subject safety is always the first priority (Figs. 2.2 and 2.3).



HEADER: PI NAME, Protocol or IRB Number, Protocol Short Title

Subject Initials ☐☐☐ **Subject ID** ☐☐☐☐☐

Serious Adverse Event **Page 1 of 2**

Event Start Date*: ☐☐ / ☐☐ / ☐☐ (Month / Day / Year)

Event End Date: ☐☐ / ☐☐ / ☐☐ (Month / Day / Year)

Date Reported*: ☐☐ / ☐☐ / ☐☐ (Month / Day / Year)

Reported to Research Staff By*: ☐☐☐☐☐☐☐☐☐☐

Death Date: *(If applicable)* ☐☐ / ☐☐ / ☐☐ (Month / Day / Year)

Death Occurred: (check one)
- ☐ Within 24 hours of investigational therapy
- ☐ Within 7 days of investigational therapy
- ☐ Within 30 days of investigational therapy
- ☐ After 30 days of investigational therapy

Did the SAE occur at your site or at a site for which the PI is responsible? ☐ Yes ☐ No

SAE Description/Narrative:

Treating Physician Comments:

PI Comments:

Outcome*: (check one)
- ☐ Fatal/Died
- ☐ Intervention for AE Continuing
- ☐ Not Recovered/Not Resolved
- ☐ Recovered/Resolved with Sequelae
- ☐ Recovered/Resolved without Sequelae
- ☐ Recovering/Resolving

Consent Form Change Required? ☐ Yes ☐ No

SAE Classification: (check all that apply)
- ☐ Fatal (resulted in death)
- ☐ A life-threatening occurrence
- ☐ Requires inpatient hospitalization or prolongation of existing hospitalization
- ☐ Results in persistent or significant disability/incapacity
- ☐ Results in congenital anomaly/birth defect
- ☐ A significant medical incident that, based upon appropriate medical judgment, may jeopardize the subject and require medical or surgical intervention to prevent one of the outcomes listed above.
- ☐ Loss of confidentiality that results in criminal or civil liability for participation or damage to financial standing, employability, insurability or reputation of the participant

Fig. 2.2 Example of severe adverse event form

HEADER: PI NAME, Protocol or IRB Number, Protocol Short Title

Subject Initials ☐☐☐ **Subject ID** ☐☐☐☐☐

Serious Adverse Event **Page 2 of 2**

SAE Reported Symptom:

Category*: [refer to the Safety Profiler website to search the Category and Toxicity of the SAE symptom reported: http://safetyprofiler-ctep.nci.nih.gov/CTC/CTC.aspx]

Toxicity*:

Grade/Severity*: (check one)
- ☐ 1 - Mild
- ☐ 2 - Moderate
- ☐ 3 - Severe
- ☐ 4 - Life Threatening
- ☐ 5 – Death (Fatal)

Unexpected*? ☐ Yes ☐ No

Dose Limiting Toxicity (DLT)? ☐ Yes ☐ No ☐ Not Applicable

Action Taken:
- ☐ Dose Reduced
- ☐ Dose Interrupted, then Reduced
- ☐ None
- ☐ Regimen Interrupted
- ☐ Therapy discontinued

Primary Attribution*: (check one)
- ☐ Definite
- ☐ Probable
- ☐ Possible
- ☐ Unlikely
- ☐ Unrelated

Detailed Attribution: (check one)
- ☐ Disease/Condition Specify:
- ☐ Investigational Treatment Specify:
- ☐ Non-investigational Treatment Specify:
- ☐ Other Specify:

Principal Investigator Signature: _____ Date: _____

Fig. 2.2 (continued)

2.6 DSMB

The FDA recommends involving a data and safety monitoring committee in the system-level management of serious adverse events [30]. A Data and Safety Monitoring Board (DSMB) is required for multisite clinical trials with interventions that involve potential risk to the participants [31]. The NIH also requires a DSMB for any NIH-supported clinical trial that has direct implications for clinical care and/or public health, involves a highly vulnerable patient population, or a high-risk intervention. The role of the DSMB is to verify that:

1. The rights and well-being of human subjects are protected.
2. The reported trial data are accurate, complete, and verifiable (esp. CRFs and AEs).

HEADER: PI NAME, Protocol or IRB Number, Protocol Short Title

Subject Initials ☐☐☐ **Subject ID#** ☐☐☐☐☐ **Page** ☐☐ **of** ☐☐

Question to ask subject: Have you had any physical, emotional or behavioral problems, symptoms, or side effects since your last visit on [DATE]? Include anything that *started* to become a problem during this time or an old problem that *worsened*.

Adverse Event Tracking Log

#	Date Reported	Adverse Event Description	Adverse Event Category**	Start Date	End Date	Ongoing (Y or N)	Outcome[1]	Severity[2]	Serious (Y or N)	Expected (Y or N)	Treatment[3]	Action Taken[4]	Attribution[5]	PI Initials

**look up corresponding AE Category at: http://safetyprofiler-ctep.nci.nih.gov/CTC/CTC.aspx

Outcome[1]
0 – Fatal
1 – Intervention continues
2 – Not recovered/not resolved
3 – Recovered w/sequelae
4 – Recovered w/o sequelae
5 – Recovering/Resolving

Severity[2]
1 – Mild
2 – Moderate
3 – Severe
4 – Life-threatening
5 – Fatal

Treatment[3]
0 – None
1 – Medication(s)
2 – Non-medication TX
3 – Subject discontinued

Action Taken[4] with Study Treatment
0 – Not Applicable
1 – None
2 – Interrupted
3 – Discontinued
4 – Dose reduced
5 – Dose increase

Attribution[5]
1 – Unrelated
2 – Unlikely
3 – Possible
4 – Probable
5 – Definite

Fig. 2.3 Example of adverse event log

The conduct of the trial is in compliance with good clinical practice, the IRB-approved protocol, and with any applicable regulatory requirement(s) [10].

2.7 CONSORT

The Consolidated Standards of Reporting Trials (CONSORT) statement, which was first written in 1996 and updated in 2010, provides guidance for reporting the aims, methods, results, and implications of randomized controlled trials. This statement includes a 25-item checklist and study flow diagram to guide the reporting and publication of an RCT [32]. At the start of the design of a clinical trial, a principal investigator is advised to review the CONSORT checklist and flow diagram and the *examples* of good reporting provided on the CONSORT website (http://www.consort-statement.org/consort-statement/). By reviewing the examples a PI can gain understanding of what methodologies and processes were used during a trial to ensure standardization and quality of the trial outcomes. In 2008, the CONSORT group provided an "extension" document for non-pharmacologic treatment interventions [33]. This document provides a 22-item checklist which is more applicable to surgical or device interventions, and the surgical trialist is advised to review this statement and its notable examples.

2.8 Conclusions

The surgical trialist is tasked with assuring that a clinical trial will be valid and generalizable to their population of interest but also needs to standardize processes and demand quality to ensure the safety of the consented subjects. This chapter provides a small introduction to the large group of acronyms that accompany the design and implementation of a clinical protocol. The academic surgical scientist interested in translational research and human clinical trials is advised to review the "regulatory and normative documents" and identify key mentors both within a Department of Surgery and within their institutions' clinical trials and statistical offices to begin to learn and speak the language associated with clinical research.

Acknowledgments Coauthor Marquita Decker is supported by an NIH T32 training grant in surgical oncology (2T32 CA090217). Tracy Ohrt, administrative program specialist at the University of Wisconsin Institute for Clinical and Translational Research, designed the case report and adverse event form templates in Figs. 2.1, 2.2, and 2.3. The forms were developed with the guidance of the CDISC Clinical Data Acquisition Standards Harmonization (CDASH) publication (http://www.cdisc.org/stuff/contentmgr/files/0/9b32bc345908ac4c31ce72b529a3d995/misc/cdash_std_1_0_2008_10_01.pdf. The UW Institute for Clinical and Translational Research is funded through a grant from the Clinical and Translational Science Award program, through the NIH National Center for Advancing Translational Sciences (UL1TR000427).

References

1. Horton R. Surgical research or comic opera: questions but few answers. Lancet. 1996;347:946.
2. Jarman AF, Wray NP, Wenner DM, Ashton CM. Trials and tribulations: the professional development of surgical trialists. Am J Surg. 2012;204:339–46.
3. Ashton CM, Wray NP, Jarman AF, Kolman JM, Wenner DM, Brody BA. A taxonomy of multinational ethical and methodological standards for clinical trials of therapeutic interventions. J Med Ethics. 2011;37:368–73.
4. Freedman B. Research, unethical. In: Reich WT, editor. Encyclopaedia of bioethics. New York: Free Press; 1995. p. 2258–61.
5. Shuster E. Fifty years later: the significance of the Nuremberg Code. N Engl J Med. 1997;337:1436–40.
6. Osler W. The evolution of the idea of experiment. Trans Congr Am Phys Surg. 1907;7:1–8.
7. World Medical Association. Declaration of Helsinki Ethical Principles for Medical Research Involving Human Subjects. Adopted by the 18th WMA General Assembly, Helsinki, Finland. 1964. http://www.wma.net/en/30publications/10policies/b3/. Accessed 1 Jan 2013
8. Lo B. Addressing ethical issues. In: Hulley SB, editor. Designing clinical research. Philadelphia: Lippincott Williams and Wilkins; 2001. p. 215–6.
9. U.S. Department of Health and Human Services, Food and Drug Administration.Using a Centralized IRB Review Process in Multicenter Clinical Trials. http://www.fda.gov/RegulatoryInformation/Guidances/ucm127004.htm. Accessed 1 Jan 2013.
10. International Conference on Harmonization of Technical Requirements of Pharmaceuticals for Human Use. ICH Harmonized Tripartite Guideline, Guideline for Good Clinical Practice E6 (R1). http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E6_R1/Step4/E6_R1__Guideline.pdf. Accessed 1 Jan 2013; ICH Harmonised Tripartite Guideline,

Structure & Content of Clinical Study Reports E3. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E3/E3_Guideline.pdf. Accessed 1 Jan 2013.
11. Code of Federal Regulations. Basic HHS Policy for Protection of Human Research Subjects. Title 45 Public Welfare, Department of Health and Human Services Part 46 Protection of Human Subjects. http://www.hhs.gov/ohrp/humansubjects/guidance/45cfr46.html. Accessed 1 Jan 2013.
12. Code of Federal Regulations. Investigational New Drug Safety Reporting Requirements for Human Drug and Biological Products and Safety Reporting Requirements for Bioavailability and Bioequivalence Studies in Humans. Title 21 Parts 312 and 320, Docket No. FDA–2000–N–0108. Washington, DC: GMP Publications; 2010.
13. Swedish Rectal Cancer Trial. Improved survival with preoperative radiotherapy in resectable rectal cancer. N Engl J Med. 1997;336:980–7.
14. Peeters KC, Marijnen CA, et al. The TME trial after a median follow-up of 6 years: increased local control but no survival benefit in irradiated patients with resectable rectal carcinoma. Ann Surg. 2007;246:693–701.
15. Richesson RL, Nadkarni P. Data standards for clinical research data collection forms: current status and challenges. J Am Med Inform Assoc. 2011;18:341–6.
16. National Cancer Institute. Case Report Forms Wiki. https://wiki.nci.nih.gov/display/CRF/Case+Report+Forms+Wiki. Accessed 1 Jan 2013.
17. Glickman SW, et al. Ethical and scientific implications of the globalization of clinical research. N Engl J Med. 2009;360:816–23.
18. Clinical Trials Transformation Initiative. Improving Unexpected SAE reporting to IND Investigators – Background. https://www.ctti-clinicaltrials.org/project-topics/adverse-events/improving-unexpected-sae-reporting-to-ind-investigators/background. Accessed 27 Dec 2012.
19. US Department of Health & Human Services, Office for Human Research Protections. Guidance on Reporting Incidents to OHRP. 2011. http://www.hhs.gov/ohrp/compliance/reports/index.html. Accessed 27 Dec 2012.
20. Medical Dictionary for Regulatory Activities Maintenance and Support Services Organization (MedDRA MSSO). http://www.meddramsso.com/public_faq_meddra.asp. Accessed 27 Dec 2012.
21. National Cancer Institute. NCI Guidelines for Investigators: Adverse Event Reporting Requirements for DCTD (CTEP AND CIP) AND DCP INDs AND IDEs. 2012. http://ctep.cancer.gov/protocolDevelopment/electronic_applications/docs/aeguidelines.pdf. Accessed 27 Dec 2012.
22. King A, et al. Time to listen: a review of methods to solicit patient reports of adverse events. Qual Saf Health Care. 2010;19:148–57.
23. Oken A, Rasmussen MD, Slagle JM, et al. A facilitated survey instrument captures significantly more anesthesia events than does traditional voluntary event reporting. Anesthesiology. 2007;107:909–22.
24. Weingart SN, Hamrick HE, Tutkus S, et al. Medication safety messages for patients via the Web portal: the MedCheck intervention. Int J Med Inform. 2008;77:161–8.
25. U.S. Department of Health and Human Services, Food and Drug Administration. The Patient Reported Outcomes (PRO) Consortium. http://www.fda.gov/AboutFDA/PartnershipsCollaborations/PublicPrivatePartnershipProgram/ucm231129.htm. Accessed 27 Dec 2012.
26. Patient Centered Outcomes Research Institute. http://www.pcori.org/. Accessed 27 Dec 2012.
27. US Department of Health & Human Services, National Institutes of Health. NIH Guidelines for Research Involving Recombinant DNA Molecules (NIH Guidelines). 2011. http://oba.od.nih.gov/rdna/nih_guidelines_oba.html. Accessed 27 Dec 2012.
28. Virzi L, Pemberton V, Ohye RG, et al. Reporting adverse events in a surgical trial for complex congenital heart disease: the Pediatric Heart Network Experience. J Thorac Cardiovasc Surg. 2011;142:531–7.

29. Wilsher CS. A qualitative examination of FDA warning letters – what can we learn from GCP inspections? Qual Assur J. 2002;6:143–57.
30. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Biologics Evaluation and Research (CBER), Center for Drug Evaluation and Research (CDER), Center for Devices and Radiological Health (CDRH). Guidance for Clinical Trial Sponsors Establishment and Operation of Clinical Trial Data Monitoring Committees. 2006. http://www.fda.gov/downloads/RegulatoryInformation/Guidances/ucm127073.pdf. Accessed 27 Dec 2012.
31. National Heart, Lung, and Blood Institute; National Institutes of Health. NHLBI Policy for Data and Safety Monitoring of Extramural Clinical Studies. http://www.nhlbi.nih.gov/funding/policies/dsmpolicy.htm. Accessed 1 Jan 2013.
32. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. Ann Intern Med. 2010;152(11):726–32. http://www.consort-statement.org/. Accessed 1 Jan 2013.
33. Boutron I, Moher D, Altman DG, Schulz K, Ravaud P. CONSORT group. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. Ann Intern Med. 2008;148(4):295–309.

Chapter 3
Statistics: Setting the Stage

Sandra L. Wong

3.1 Introduction

Results from clinical trials are often cited as the highest-quality evidence available to guide clinical practice. While many reported clinical trials are of high quality, many have significant deficiencies in study design, data analysis, and interpretation of results. Up-front attention to the fundamentals of statistics is paramount to the design and conduct of a high-quality clinical trial. This chapter will address the fundamentals of study design, focusing on statistical topics including randomization, treatment allocation, cohort stratification, sample size, power, and type II errors. Issues directly related to data analysis are covered in another chapter and are not discussed here.

3.2 Setting the Stage for Clinical Trials

The development of a clinical trial must, by definition, include consideration of important clinical questions. Statistical considerations are an essential part of study design and are critically important to all aspects of the trial. Clinical trials are resource intensive and take a long time to conduct. At the end of the trial, nothing is more disappointing than having less than ideal results because the initial study planning was flawed. In fact, the use of the CONSORT statement [1] is now standard practice when reporting results from randomized clinical trials (RCTs), and this checklist and flowchart helps assure that trials were designed and conducted in a high-quality manner. Importantly, these standards provide uniform information about why and how a study was conducted, providing sufficient transparency to

S.L. Wong, MD, MS
Department of Surgery, University of Michigan, Ann Arbor, MI, USA
e-mail: wongsl@med.umich.edu

T.M. Pawlik, J.A. Sosa (eds.), *Success in Academic Surgery: Clinical Trials*,
Success in Academic Surgery, DOI 10.1007/978-1-4471-4679-7_3,

allow appraisal of trial results and downstream comparisons of effectiveness [2]. Critical analyses of negative results of RCTs published in the surgical literature have shown that many reports simply lack sufficient statistical power to definitively support or refute the null hypothesis being examined [3].

The statistical aspects of design must include structure of the study, collection of data, and choice of measurements to make. Careful study design occurs at the outset and is not an afterthought.

3.3 Study Design

Clinical trials are experimental and involve some type of intervention to the subjects. Not all clinical trials can or should be large in scope, and not all are randomized in design. In the surgical sciences, randomization may not be possible since a decision to have an operation or not have an operation may be a decision that a patient does not want to have made by a randomization process. Having controls in a study lends credence to inferences of causality when there is an observed association.

The RCT is the standard by which all other trial designs are judged since other designs have features that introduce bias. The process of randomization tends to produce study groups that are comparable with respect to known and unknown variables, thereby removing biases between groups. However, even with randomization, biases can remain and care must be taken to minimize biases when possible [4].

3.4 Nonrandomized Studies

Given the variability of response to an intervention and its impact on the natural history of disease, there is a need for a well-defined control or comparison group in a well-planned clinical trial. Controlled clinical trials can be conducted without randomization. There are many types of nonrandomized clinical trials, including historical control studies in which a group of participants receive an intervention and are compared to a previous control group. While an argument has been made in favor of historical control design because it allows all subjects to receive the intervention, there are significant limitations and inherent biases. Simply, patients with any given condition today are different than patients with that condition a decade or two ago. The fallacy of post hoc ergo propter hoc reasoning is steeped in seeing improvements in outcomes because the patient population or other parts of patient management have changed over time, yielding patients with earlier diagnoses or better prognoses in the treatment group. For example, there have been impressive declines in mortality from coronary artery disease over time, and any trial employing historical controls would face the nearly impossible task of separating treatment effect from secular trends. Further, because the groups are not contemporaneous,

those receiving the intervention tend to be less seriously ill, and the comparison does not adequately account for the differences in expected and observed mortality. Not surprisingly, when studies of the same therapies from randomized controlled trials were compared to historic control trials, the latter consistently yielded more favorable results [5].

Contemporary, or concurrent, controlled trials include studies with a crossover design, in which participants are used twice, once as a member of the control group and once as a member of the intervention group. The major appeal of the crossover design is the "within-subject," or paired, comparisons or an assessment of whether a single participant does better with the control or the intervention. However, the strict assumption that what happens in the first period does not carry over into the second period of the trial must be valid for this study design to even be considered, even if the participant is blinded to control and intervention. A washout period may be introduced between treatment periods, but carryover effects may still be present. Any type of carryover effect would decrease the power of the study and necessitate a larger total sample size for the study. Another concern is the period effect, since there may be some systematic difference between the first and second periods of the study.

Withdrawal studies assess the response to the discontinuation of an intervention. These types of studies may be used to evaluate the duration of benefit with an intervention and may be particularly useful when an ongoing intervention (e.g., medication) has never conclusively been shown to be beneficial. Withdrawal studies can include a randomization component, and this strategy may reduce the inherent bias that accompanies the selection of participants, typically ones with excellent results or ones with stable disease, to even be considered for withdrawal of therapy.

A factorial design evaluates two interventions, X and Y, with control in a single experiment. A commonly used factorial design is the "2×2 design" which includes four "cells": X+Y, X+control, Y+control, and control alone. For example, the Canadian transient ischemic attack study included aspirin and sulfinpyrazone, compared to placebo [6]. The most important assumption for factorial design studies is that there is no interaction. For instance, the effect of intervention X may differ based on the presence or absence of Y. In the case of the transient ischemic attack study, interaction was a major concern because the drugs may have acted through the same mechanism of action [7].

Group allocation designs, or cluster randomization studies, have the appeal of randomizing at the hospital (group) level, thereby eliminating the need to randomize at the participant level; all participants at a certain center are controls or receive a specific intervention. The effective sample size and units of analysis, however, consists of the groups (not total number of participants) and limits the efficiency of these studies. Historically, these groups have been communities participating in prevention trials, but cluster randomization studies have gained some acceptance among surgeons evaluating different surgical approaches or techniques. However, individual participants are grouped within each hospital and cannot be regarded as statistically independent, and standard sample size calculations would severely underestimate the total number of participants [7].

3.5 Randomization Schema

In the event that an RCT is planned, randomization techniques should be explored in order to maximize enrollment while minimizing residual biases from the process of randomization itself. Randomization simply refers to a process by which all participants are equally likely to be assigned to either the intervention or control group. The randomization schema should be practical but minimize selection bias, serving three main purposes.

First, randomization has the advantage of removing bias in the allocation of a particular participant to the intervention or control group. Any undue influence in assignment can invalidate the comparison. Further, the randomization process tends to produce comparable groups. In the end, measured and unmeasured prognostic factors of the participants are balanced between the intervention and control groups. Finally, randomization allows a probability distribution which ensures the validity of statistical tests of significance used to detect observed differences between groups.

There are many randomization processes [7], ranging from very simple schemes (e.g., coin toss) to much more complex strategies. Selection bias is avoided with randomization, and randomization must not occur before a participant decides to enter a study to ensure that all the benefits of randomization are intact. Fixed allocation randomization processes make assignments to the intervention group at a predefined probability. Most commonly, there is an equal allocation to either group (1:1 allocation results in an equal probability of being assigned to intervention versus control). In some instances, less information may be needed about effects to the control group, and unequal allocations (e.g., 2:1 allocation to intervention compared to control) may be used to gain more information about the intervention. There is a loss of power with this strategy, and some trialists have concerns that unequal allocation designs are not consistent with true equipoise and may bias participants, especially if they are randomized to the control.

Fixed allocation randomization processes most commonly include simple, blocked, and stratified designs. Simple randomization employs a coin toss or random digit table to make the assignments. Alternating assignments (e.g., ABABAB) are not truly random and should be avoided. Assignment made by birthdate (e.g., odd dates vs. even dates) appears to yield unbiased allocation, but any potential knowledge of which intervention a potential participant may receive can affect decisions about whether or not to enroll. In situations with more than one group, intervals are used to make assignments (e.g., 0–0.249 is assigned to group A and 0.25–0.4999 to group B). In the long run, the number of participants in each group will be balanced, but there is a risk of imbalanced enrollment during accrual, limiting the power of the study if the trial is stopped early. A blocked randomization schema can minimize large imbalances between groups and maximize power. Using set block sizes, there would be an equal probability of assignment to each group. For example, if a block size of four is used, then the process would randomize each of four participants to group A or B, for every consecutive group of four participants.

Since many covariates are strongly associated, imbalances in one could lead to imbalances in others. Some trials may exclude participants with certain characteristics in order to maintain a homogeneous study population, but highly stringent exclusion criteria can limit enrollment and call the applicability or generalizability of the study's results into question. In these situations, stratified randomization can adjust for the imbalances by measuring the covariates of interest (e.g., smoking history) prior to or at the time of randomization. The randomization is then performed within defined stratum or strata (e.g., current smoker, former smoker, never smoker), and at that level, the randomization can then either be simple or blocked. Using this strategy, post hoc changes in the analytic plan can be avoided. In large multicenter trials, consideration should be given to stratified randomization by the center.

Adaptive randomization processes are meant to change allocation probabilities as the study progresses. Baseline adaptive randomization processes can be used to balance the number of participants in groups without consideration of treatment response or other outcomes. The response adaptive randomization processes adjust allocations according to intervention responses. Two examples of the response adaptive randomization models [7] include the "play the winner" and "two-armed bandit" approaches and rely upon relatively quick response measurements in order to make subsequent group assignments.

The "play the winner" approach bases the assignment of the second (and subsequent) participant on the prior participant's response to the intervention, making the same group assignment if the first participant had a successful response. The group assignments continue in this manner until a failure occurs, and then the assignment switches. The "two-armed bandit" method continuously updates the probability of success as the outcome for each participant becomes known, with the goal of assigning a higher proportion of future participants to the "better" intervention. While these strategies maximize the number of participants to the superior intervention, there is a loss of power associated with the resulting imbalances in the groups.

Large simple clinical trials, not to be confused with pragmatic clinical trials, are based on a model of relatively straightforward interventions and easily ascertained and measured outcomes. These types of trials are best suited for common conditions (to allow for large numbers of participants) in which even modest benefits of an intervention would be important to uncover. The GUSTO trial which evaluated four thrombolytic treatment strategies for acute myocardial infarction is an excellent example of a so-called large-scale trial which enrolled over 41,000 patients at over 1,000 hospitals [8].

However, RCTs are not always as straightforward as deciding on a control and an intervention. Some investigators, even in the setting of equipoise, have ethical conflicts about "depriving" a participant of a new, and presumed improved, intervention and may avoid trial participation because of these conflicts. Some trials may fail to accrue adequate numbers of participants because many drop out if they are not randomized to the "preferred" intervention group [9].

Some unique features of surgical trials warrant special consideration. A proposed taxonomy for surgical trials includes three broad categories: comparisons of minor variations in operative techniques, comparisons of major variations in

technique, and comparisons of surgical and nonsurgical treatments [10]. With all of these, intrinsic technical variability must be considered and accounted for whenever possible. Standardization and reproducibility are much more difficult to control when specific nuances of procedures are being examined. One excellent example of this issue is a trial of extended lymph node dissection in gastric cancer which was conducted by the Dutch Gastric Cancer Group [11]. The goal of the trial was to examine whether extended (D2) lymph node dissection with gastrectomy improved survival and reduced recurrent disease compared to a more limited (D1) dissection. The trial and its results faced criticism because, even with strict attempts to teach and standardize the procedure, there were significant deviations from the surgical protocol attributed to relatively low-volume surgeons and difficulty achieving a true D2 dissection when that was the intent [12].

Major variations in technique can also be compared, and clinical trials are often important in establishing them as viable options. Recent examples include the trial comparing a laparoscopic to open approach for colon resection by the Clinical Outcomes of Surgical Therapy (COST) Study Group [13]. Comparisons like this may also encompass other "competing" treatment modalities. For example, major trials have evaluated percutaneous interventions for coronary artery disease to open surgical approaches [14].

Of course, when the effectiveness of a specific procedure is compared to nonoperative therapy, technical considerations may be somewhat less of a consideration, as with a Gynecologic Oncology Group trial comparing secondary cytoreduction surgery for stage III ovarian cancer after chemotherapy to no additional cytoreduction [15]. Direct comparisons of a surgical intervention to medical management have been done as well, such as a comparison of laparoscopic fundoplication to proton pump inhibitors [16] or surgery versus nonsurgical therapy for carpal tunnel syndrome [17] or back pain [18].However, the admonition that the quality of a surgical procedure can even impact clinical trials evaluating medical therapies is well heeded since differences in the extent of resection and downstream impact on disease control can mask small effects of systemic treatments [19].

Surgeons have also taken the lead on trials that do not necessarily evaluate a procedure. Because surgeons are the first line of care for many diseases, surgeons are also the first line of referral for subsequent treatments. With the development of a surgically oriented cooperative trial group by the National Cancer Institute, the American College of Surgeons Oncology Group (ACOSOG), now part of the Alliance, was established in 1998. One of the most successful ACOSOG trials was a randomized double-blind, placebo-controlled trial testing the adjuvant use of imatinib after resection of gastrointestinal stromal tumor [20]. Another example of a "nonsurgical" trial is evaluating postoperative feeding regimens in a pediatric population [21].

The recent emphasis on pragmatic clinical trials (PCTs) warrants separate consideration. The goal of these PCTs is to evaluate interventions in real-world, community settings, the results of which are meant to be distinguished from standard RCTs which are conducted under optimal settings in carefully chosen patients [22]. Generally, PCTs would have less inclusion and exclusion criteria and enroll participants at a broader range of study sites as opposed to only academic medical centers.

Comparative effectiveness research (CER) does not classically include actual clinical trials, but results from clinical trials are used as the evidence base for CER. That is, existing evidence from clinical trials is combined with data from other studies in CER to inform clinical decision making when data from clinical trials is not sufficient in itself. For example, key questions such as how open hernia repair differs from laparoscopic hernia repair were addressed in a recent comparative effectiveness study [23].

3.6 Endpoints/Response Variables

As the study is planned, careful consideration must be given to primary and secondary endpoints of the trial [7, 24]. Primary response variables are used to judge the effectiveness of the intervention (i) compared to the control (c) and may be measured during and at the end of the trial. Common response variables include a clinical finding or an adverse event. There are three basic kinds of response variables: dichotomous, continuous, and time to event. Dichotomous response variables are typically event rates (p) for the response of interest. For continuous response variables, the mean level (μ) of response is compared between groups. For time to event measurements, the hazard rate (λ) is compared between groups.

When there are several response variables, such as multiple endpoints of interest or response variables that are repeatedly monitored over time, a combination of response variations may be used, typically requiring a defined hierarchy of events of interest. There may be other endpoints of interest, but analysis of those findings is secondary to the primary intent of the study, and results must be interpreted with caution since up-front planning and sample size calculations are based on the primary endpoints only.

Surrogate response variables such as biologic endpoints may need to be considered in clinical trials. In general, in order for a surrogate response variable to have meaning, it must capture the relationship between the treatment and the corresponding true endpoint [25]. In other words, the biologic mechanism or pathway must be known. Use of these variables has the advantage of shorter follow-up time needed and lower resource utilization (ease of data collection, cost-effectiveness). The surrogate fails if it is insensitive to the effect of the treatment on the true outcome of interest or if it is only representative of one clinical endpoint, but not others.

3.7 Statistical Power [7]

Results of clinical trials are used to make inferences. Sound biologic inferences are guided by consistent, reproducible observations, which in turn are tested with statistics. An important, and sometimes underappreciated, component of trial design is the up-front statistical planning which dictates how the study is conducted and how the results will be interpreted. The trade-offs between sample size, power, statistical

significance, and detectable differences between the groups must be carefully weighed a priori, especially in the context of conducting a feasible clinical trial in terms of ability to accrue patients in a timely manner.

Typically, the null hypothesis (H_o), which states that there is no difference between event rates (H_o: $p_c - p_i = 0$), mean levels (H_o: $\mu_c - \mu_i = 0$), and hazard rates (H_o: $\lambda_c - \lambda_i = 0$), is tested. Even if H_o is true, the observed response variables may by chance show a difference. If observed differences are large enough by chance alone and H_o is then incorrectly rejected, this false-positive occurrence is termed a type I error (α). The probability of a type I error is called the significance level and is known as the "p value (p)." That is, it is the probability of observing differences as large as the difference actually observed given that H_o is true (e.g., there is no difference) is α. The decision to reject H_o is typically based on a level of α of 0.01 or 0.05 (e.g., "the probability that the observed differences occurred by chance when there is actually no difference is α"). Typically, the smaller the α is set at, the larger the required sample size is. The common use of p values to bolster a study's significance reflects the appropriateness of rejecting H_o.

A type II error (β) occurs when H_o is not rejected when there is actually a difference. This false-negative result is typically attributed to a small number of observed differences, especially when a larger number of observations would have had a better chance of detecting clinically meaningful differences. Indeed, power is denoted as 1-β and quantifies the ability of a study to find true differences. β is a function of α, the sample size, and the true difference between response variables (commonly denoted δ). Typically, α is set low, at 0.05 or 0.01, and β is set high, at 0.90 or 0.95, so a trialist can vary the expected difference δ to be detected and/or the total sample size when planning a clinical trial. Previous research or historic results may be used to inform estimates of δ. The selection of δ should be a level that is the minimal difference between groups that would still be clinically meaningful. Estimates of δ are somewhat artificial constructs but should be guided by the degree of benefit that the intervention would have to have in order for it to be worthwhile.

The primary purpose of sample size calculations is to ensure a high likelihood of detecting a statistically significant, clinically meaningful effect of the intervention being tested if it exists. Similarly, there should be an assurance that no such benefit exists if the trial does not demonstrate one. The plot of δ (horizontal axis) versus 1-β (vertical axis) for any given sample size and set α is the power curve (Fig. 3.1) and demonstrates the probability (or power) of detecting a given true difference. In general, if the treatment effect is very large, a smaller sample size could be used with adequate power. However, most trials that are conducted are unpowered, probably reflecting lack of attention to adequate sample size calculations during the planning phase.

3.8 Sample Size Considerations

Clinical trials should have sufficient statistical power to detect clinically meaningful differences between groups. Calculation of the necessary sample size for adequate levels of significance and power is an essential part of the planning process for

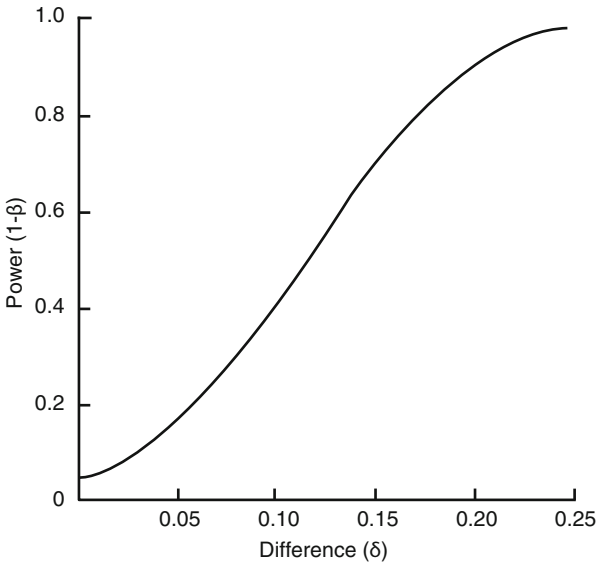


Fig. 3.1 Power curve. A power curve demonstrates the increasing differences (δ) between the control group and the intervention group with a one-sided significance level of 0.05 and a total sample size ($2N$) (From: *Sample Size*. Friedman et al. [7]. Figure 7.1, page 97)

clinical trials. Sample size estimates are based on multiple assumptions and therefore should be as conservative as can be justified. Overestimated sample sizes may lead to unfeasible enrollment goals, but underestimates may lead to incorrect conclusions at the end of the trial.

To delineate how sample size calculations are made, the following example will be based upon a trial comparing one intervention group with one control group. If randomization results in equal allocation to the intervention and control groups, an assumption that the variability in responses between the groups is similar can be made, and the sample size can be equally split (in the case of a 1:1 randomization schema).

Differences can be tested as a one-sided test or a two-sided test. In other words, the trialist must decide if the study is going to measure differences in one direction (i.e., there are improvements in the intervention over the control) or differences in either direction (i.e., the intervention is either better or worse than the control). Because a new intervention could be either beneficial or harmful, two-sided tests are preferred. In the unusual circumstance in which a difference is only expected in one direction, the significance level used in calculating the sample size should be half of what would be used for a two-sided test.

Interim analyses may be employed in a clinical trial. If early data suggest that an intervention is harmful, the trial may have to be terminated prematurely for safety reasons. Further, if results appear unlikely to show any difference at the end of the trial, early termination may be considered to spare the resources necessary for trial continuation. Conversely, if early data suggest that an intervention is clearly beneficial, the trial may similarly have to be terminated prematurely because to continue with a control group may be unethical. One specific consequence of early data analysis is that the rate of incorrectly rejecting H_o will be larger than the initially selected significance level unless the critical value for an interim analysis is appropriately adjusted. Importantly, the more interim analyses, the higher the probability of incurring a type I error.

Equipoise is the guiding principle behind RCTs; without true uncertainty about an intervention's superiority, there is no justification for conducting a clinical trial. Sometimes, the appropriate trial is a study of equivalency, also known as trials with positive controls. Equivalency studies test whether a new intervention is as good as an established one. The study design here includes a standard intervention (control) which is known to be better than placebo. The new intervention may be preferred because it is less expensive, has fewer side effects or other attractive features. Importantly, no trial can ever show that the two interventions are the same; there is no way to demonstrate complete equivalence ($\delta = 0$) because this would require an infinite sample size. However, it is reasonable to specify a value of δ such that a difference is less than might be considered equally effective or equivalent. Therefore, the null hypothesis states that differences are greater than δ and failure to reject H_o merely finds that there is inadequate evidence to say that the groups are different.

3.9 Conclusion

Proper attention to statistics when planning and conducting a clinical trial will help prevent an effective therapy from being disregarded because of false-negative results or a relatively ineffective therapy from applied improperly because of false-positive results.

References

1. Begg C, Cho M, Eastwood S, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. JAMA. 1996;276:637–9.
2. Tiruvoipati R, Balasubramanian SP, Atturu G, Peek GJ, Elbourne D. Improving the quality of reporting randomized controlled trials in cardiothoracic surgery. The way forward. J Thorac Cardiovasc Surg. 2006;132:233–40.
3. Dimick JB, Diener-West M, Lipsett PA. Negative results of randomized clinical trials published in the surgical literature. Equivalency or error? Arch Surg. 2001;136:796–800.
4. Designing research. In: Altman DG, editor. Practical statistics for medical research. London: Chapman & Hall/CRC; 1999.
5. Sacks MH. The interaction between clinical research and treatment. Hillside J Clin Psychiatry. 1983;5:169–76.
6. The Canadian Cooperative Study Group. A randomized trial of aspirin and sulfinpyrazone in threatened stroke. N Engl J Med. 1978;229:53–9.
7. Friedman LM, Furberg CD, DeMets DL. Fundamentals of clinical trials. New York: Springer; 1998.
8. The GUSTO Investigators. An international randomized trial comparing four thrombolytic strategies for acute myocardial infarction. N Engl J Med. 1993;329:673–82.
9. Randomized clinical trials. In: Levine RJ, editor. Ethics and regulations of clinical research. 2nd ed. New Haven: Yale University Press; 1988.
10. Blazeby JM, Donovan J, Sharples L, Morton D. Recruiting patients into randomized clinical trials in surgery. Br J Surg. 2012;99:307–8.

11. Bonenkamp JJ, Hermans J, Sasako M, et al. Extended lymph-node dissection for gastric cancer. N Engl J Med. 1999;340:908–14.
12. Brennan MF. Lymph-node dissection for gastric cancer. N Engl J Med. 1999;340:956–8.
13. Fleshman J, Sargent DJ, Green E, et al. Laparoscopic colectomy for cancer is not inferior to open surgery based on 5-year data from the COST Study Group trial. Ann Surg. 2007;246:655–62.
14. Serruys PW, Morice MC, Kappetein AP, SYNTAX Investigators, et al. Percutaneous coronary intervention versus coronary-artery bypass grafting for severe coronary artery disease. N Engl J Med. 2009;360:961–72.
15. Rose PG, Nerenstone S, Brady MF, et al. Secondary surgical cytoreduction for advanced ovarian carcinoma. N Engl J Med. 2004;351:2489–97.
16. Grant AM, Wileman SM, Ramsay CR, et al. Minimal access surgery compared with medical management for chronic gastro-oesophageal reflux disease: UK collaborative randomized trial. BMJ. 2008;337:a2664.
17. Jarvik JG, Comstock BA, Kliot M, et al. Surgery versus non-surgical therapy for carpal tunnel syndrome: a randomized parallel-group trial. Lancet. 2009;374:1074–81.
18. Birkmeyer NJO, Weinstein JN, Tosteson ANA, et al. Design of the spine patient outcomes research trial (SPORT). Spine. 2002;27:1361–472.
19. Reynolds T. Why randomized surgical oncology trials are so scarce. J Natl Cancer Inst. 1999;91:1182–3.
20. DeMatteo RP, Ballman KV, Antonescu CR, et al. Adjuvant imatinib mesylate after resection of localized, primary gastrointestinal stromal tumour: a randomized, double-blind, placebo-controlled trial. Lancet. 2009;373:1097–104.
21. Ong EGP, Eaton S, Wade AM, et al. Randomized clinical trial of glutamine-supplemented versus standard parenteral nutrition in infants with surgical gastrointestinal disease. Br J Surg. 2012;99:929–38.
22. Tunis SR, Stryer DB, Clancy CM. Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy. JAMA. 2003;290:1624–32.
23. Evidence-based practice center systematic review protocol. Comparative Effectiveness of Surgical Options for Inguinal Hernia. Accessed at: www.effectivehealthcare.ahrq.govpublished. 15 Aug 2011.
24. Clinical trials. In: Altman DG, editor. Practical statistics for medical research. London: Chapman & Hall/CRC; 1999.
25. Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. Stat Med. 1989;8:431–40.

Chapter 4
Clinical Trials: Handling the Data

Christina E. Bailey and George J. Chang

4.1 Introduction

Randomized control trials (RCTs) provide the foundation for evidence-based medicine, which is the cornerstone of medical practice. RCTs are prospective studies that compare the effect of an intervention between an intervention and control group. An understanding of statistical methods is fundamental to the interpretation of RCT methods and results. This chapter will not provide an in-depth description of the methods of statistical analysis (this information can be obtained from any introductory statistics textbook). Instead, this chapter will provide a brief review of common statistical methods used to analyze data and discuss some issues associated with data analysis.

4.2 Who Should Be Analyzed

The first question that should be answered before proceeding with data analysis is which study participants should be included in data analysis. Defining the study population has important implications for the feasibility of the study and generalizability of the results. Unfortunately, even some of the best-designed clinical trials often cannot be perfectly implemented. In retrospect, some participants may not have met the inclusion criteria, data for some participants may be missing, or the protocol may not have been completely followed. Some investigators prefer to eliminate participants who do not adhere to the inclusion criteria or the protocol, whereas other investigators believe that once a participant is randomized, he or she

C.E. Bailey, MD, MSCI • G.J. Chang, MD, MS (✉)
Department of Surgical Oncology, University of Texas, MD Anderson Cancer Center,
1400 Pressler Street, FCT 17.6060, 301402, Houston, TX 77230-1402, USA
e-mail: gchang@mdanderson.org

T.M. Pawlik, J.A. Sosa (eds.), *Success in Academic Surgery: Clinical Trials*,
Success in Academic Surgery, DOI 10.1007/978-1-4471-4679-7_4,

should be included in the final analysis [1]. Both of these views will be discussed later.

Exclusions are potential participants who do not meet all of the entry requirements and are not randomized. Fortunately, exclusions do not bias the results, but it is important to document exclusion criteria in the trial protocol because exclusions can influence interpretation of the results. Participants may also be withdrawn from analysis. Multiple reasons exist for withdrawing participants from the analysis, including ineligibility, nonadherence, and poor quality and/or missing data. Participant withdrawal can bias the results, and it is important to develop a policy on the handling of withdrawals during the design of the trial. Investigators are responsible for convincing readers that the analysis was not biased secondary to participant withdrawal.

4.3 Expressing the Data

Before reviewing common statistical methods used to analyze data, we will first review hypothesis testing. Hypothesis testing allows investigators to make generalizations from a sample to the population from which the sample was obtained [2]. The first step in hypothesis testing is stating a null (H_0) and alternative (H_A) hypothesis. The H_0 states that there is no difference between the hypothesized and population mean, whereas the H_A states that there is a difference between the hypothesized and population mean. The next step in hypothesis testing is to decide on the appropriate statistical test (reviewed in greater detail below). It is essential to account for random variation in order to conclude that the observed differences in the samples are not due to chance. The p-value estimates the probability of a true difference occurring by chance. If the observed results are highly unlikely (i.e., $p<0.05$), we reject the H_0 and accept the H_A. This means that 5 times out of 100, we will reject the H_0 when it is true (i.e., state there is a difference between two populations when a difference does not exist). This is referred to as type 1 or alpha (α) error. Conversely, type 2 or beta (β) error occurs when an investigator fails to reject the H_0 when it is false (i.e., state there is no difference between populations when a difference does exist). Power (1-β) is the ability of a study to detect a true difference, and is important in hypothesis testing. Whereas α error of 0.05 is conventionally accepted, β error of 0.10 or 0.20 is most often used (i.e., power of 90 or 80 %).

An example of utilizing hypothesis testing is as follows. Researchers implemented a trial to determine whether work as a fire fighter affects pulmonary function tests. The study included 50 firefighters, and forced expiratory volume after 1 second (FEV_1) was measured before and after a 2-year period on the job. The expected mean decline in FEV_1 over 5 years in normal males is 0.10 L. In this study, the H_0 is the mean decline in FEV_1 will be equal to 0.10 L, and the H_A is the mean decline in FEV_1 will not be equal to 0.10 L.

4.3.1 Comparison of Two Means

The student's *t*-test can be used to determine whether the means of two separate groups are equal. Student's *t*-test compares the means of two continuous variables and expresses the probability that any differences are due to chance or a "real" difference exists. Data can be obtained from paired or unpaired samples. Paired samples occur when observation are made in the same person (involves a before and after treatment measurement), and unpaired samples occur when observations in one group are independent from observations in another group [3]. There are three assumptions that must be met to utilize the student's *t*-test: the data in both groups must follow a normal distribution, the standard deviation (or variance) for both groups is equal, and both groups are independent. Violation of any of these three assumptions can lead to misleading conclusions. If the assumptions are violated, it is recommended that a nonparametric method (Mann–Whitney *U* test for unpaired data or Wilcoxon signed rank test for paired data) be used instead. The one-way analysis of variance (ANOVA) is used to compare the means of three or more groups.

An example of the student's *t*-test can be illustrated using the previous example comparing FEV_1 among firefighters before and after a 5-year period on the job. As mentioned earlier, the expected mean decline in FEV_1 over 5 years in normal males is 0.10 L. The mean decline in FEV_1 in the 50 firefighters included in the study is 0.2 L, and using student's *t*-test to compare means gives a p-value < 0.001. One therefore rejects the H_0 and concludes that the observed decline in FEV_1 is significantly different from the expected decline.

4.3.2 Comparison of Two Proportions

The chi-square test and Fisher's exact test can be used to compare frequencies or proportions in two or more groups [4]. For example, consider a clinical trial comparing a new treatment (Drug A) to reduce mortality after pulmonary embolus. The primary end point is survival or death. A total of 1,000 patients with pulmonary embolus were randomized to receive Drug A ($n = 525$) or placebo ($n = 575$). In the treatment group 27 patients (5 %) died and in the placebo group 75 (13 %) died. The data can be displayed in a 2×2 table.

	Died	Survived	
Drug A	27	498	525
Placebo	75	500	575
	102	998	1,100

The row totals are the total number of patients receiving Drug A and placebo, whereas the column totals are the total number of patients who died and survived. The chi-square test can be used to determine if there is a statistically significant association between death and treatment with Drug A. The H_0 would be there is no association between death and treatment with Drug A, and the H_A would be there is an association between death and treatment with Drug A. Using the chi-square test, $p < 0.001$ therefore rejecting the H_0, and there is an association between death (or improved survival) with Drug A. Of note, the Fisher's exact test is used when the expected cell frequencies are <5. The expected cell frequency is the probability of being in a given cell times the total sample size. For example, the expected cell frequency for the upper left cell is calculated as $(525 \times 102)/1{,}100 = 48.7$.

4.3.3 Relative Risk and Odds Ratio

The relative risk (RR) is the ratio of the incidence in people with the risk factor (exposed persons) to the incidence in people without the risk factor (nonexposed persons). RR can only be calculated for cohort studies and clinical trials. In both instances, there is a group of subjects with the risk factor and a group of subjects without the risk factor. The subjects are then followed over time to determine which subjects develop the outcome of interest.

The odds ratio (OR) is the odds that a subject with an adverse event was at risk divided by the odds that a subject without an adverse event was at risk. OR can be calculated for cohort and case–control studies. The OR and RR can be easily calculated using a 2×2 table.

	Disease	No disease
Treated/exposed	a	b
Control group	c	d
$OR = a * d/b * c$		
$RR = \frac{a/(a+b)}{c/(c+d)}$		

For example, a trial was performed comparing thrombotic events in patients taking a nonsteroidal anti-inflammatory drug (NSAID) compared to placebo. In the NSAID group, 46 out of 1,000 patients had a thrombotic event compared to 26 out of 1,000 patients in the placebo group.

	Thrombotic event		Total
	Yes	No	
NSAID	46	954	1,000
Placebo	26	974	1,000

The calculated OR would be 1.81 [(46 × 974)/(954 × 26)]. This can be interpreted as patients taking the assigned NSAID have 1.81 increased odds of having a thrombotic event compared to patients taking placebo. The calculated RR [(46/1,000)/(26/1,000)] is 1.76. This can be interpreted as patients taking the assigned NSAID have a 76 % increase in the rate of thrombotic events compared to patients taking placebo.

OR or RR greater than 1 indicates that there is an increased risk of the measured event associated with the exposure. When the OR or RR equals 1, the measured event is no more likely to occur with or without the exposure. On the other hand, when the OR or RR is less than 1, the measured event is less likely to occur with the exposure [5]. Also of note, in the previous example, OR and RR approximate each other, 1.81 and 1.76, respectively. This is usually true when the event rates are low and/or the treatment effect is small.

Other terms to be familiar with include absolute risk reduction, number needed to treat, absolute risk increase, and relative risk reduction. The absolute risk reduction allows one to assess the reduction in risk compared with the baseline risk. Specifically, it is the reduction in risk of a new intervention compared to the risk without intervention, and it is the absolute value of the difference between the experimental and control event rates. The number needed to treat is the reciprocal of the absolute risk reduction and provides the number needed to treat in order to prevent one event. For example, if a new treatment decreases the relative risk of myocardial infarction and has an absolute risk reduction of 0.0086, then the number of people who need to be treated to prevent 1 myocardial infarction is approximately 116 (1/0.0086 = 116.3). Absolute risk increase is the opposite of the absolute risk reduction. It is the increase in risk with a new treatment compared with the risk without the treatment, and relative risk reduction is the reduction in risk with a new treatment relative to the risk without treatment [4].

4.3.4 Correlation and Linear Regression

Correlation is used to determine if a linear relationship exists between two quantitative variables. Linear correlation is a measure of the degree to which an increase or decrease in one continuous variable is associated with a proportional increase or decrease in a second continuous variable [6]. In other words, can the relationship between two variables be described by a straight line? For example, consider a

scatterplot depicting the hemoglobin A1c and serum glucose in ten patients with diabetes mellitus. If every point falls on a straight line, the two variables are perfectly correlated. The Pearson correlation coefficient (r) can be used to calculate the strength of a relationship and ranges from −1 to +1. A value of 0 represents no correlation, −1 represents perfect negative correlation, and +1 represents perfect positive correlation between two variables. The Pearson's correlation coefficient can be calculated for any dataset, but it is more meaningful if the two variables are normally distributed.

Linear regression allows investigators to analyze the relationship between two or more continuous variables when one variable depends on the others and allows investigators to predict one variable given the value of the other variables [3]. For example, investigators were interested in the relationship between height and forced expiratory volume (FEV) in children. Using linear regression, it was found that FEV = −6.07 + (0.14 × height). Using this equation, the predicted FEV for a five foot (60″) child would be 2.34 l (−6.06 + 0.14 × 60). Of note, when performing multivariate analysis (i.e., more than two variables are included in the model), the number of covariates used in the model depends on the sample size. Ideally the sample size should exceed ten times the number of independent variables. For example, if the sample size in a study is 100, no more than ten independent variables should be included in the linear regression model. If too many independent variables are included in the model, investigators run the risk of overfitting the data. The same is also true for small sample size. Also, assumptions must be met in order to utilize linear regression models. They are as follows: the sample must be randomly selected, X and Y are normally distributed, and the Y values are independent of each other (i.e., not correlated).

4.3.5 Survival Analysis

Survival analysis is also referred to as time to event analysis. It allows for the analysis of binary categorical outcomes such as death, onset of disease, recurrence of disease, and onset of disability. Survival can be reported as a percentage (i.e., 1-year or 5-year survival), median survival, or survival curves. There are several different survival analysis methods: incidence density method, life table (actuarial method), Kaplan-Meier (product-limit method), and Cox Proportional Hazards model. Kaplan-Meier and Cox Proportional Hazards are the more commonly used survival methods for clinical trials. Kaplan-Meier survival analysis allows investigators to generate survival curves for each group which can be compared using the logrank statistic. Kaplan-Meier survival analysis is to be considered generally reliable up to two times the median follow-up time. Assumptions made when utilizing the Kaplan-Meier method include no change in the event rate over time and the outcome is the same for patients that are followed and those lost to follow-up. The Cox Proportional Hazard model provides a hazard ratio and allows for the comparison of two or more survival curves after adjusting for covariates. For example, a multi-institutional retrospective study identified 3,500 patients who underwent pancreatic resection for

pancreatic cancer. A multivariate-adjusted Cox Proportional Hazard model was used to evaluate the prognostic significance of adjuvant radiation therapy (AXRT). The hazard ratio for patients who received AXRT was 0.75. This can be interpreted as patient who received AXRT after surgical resection of pancreatic cancer had a 25 % decreased risk of death compared to patients who did not receive AXRT. During covariate adjustment, in general the ratio of the number of independent variables used in the Cox model to number of events should not exceed 1:10. For example, if a study has a sample size of 1,000 and 100 patients died, the maximum number of independent variables that should be included in the model is 10. Similar to linear regression, if too many independent variables are included in the model, one is at risk of overfitting the data. Assumptions for Cox regression are the same as Kaplan-Meier survival analysis, and the effect of the covariate does not change over time for any of the independent variables.

4.4 Analyzing the Data

Careful analysis of data obtained from clinical trials requires a major investment of time and effort. Inappropriate statistical analysis can result in misleading conclusions and impairs the credibility of the trial and investigators. Two important issues that should be considered in the analysis of clinical trial results are intention-to-treat analysis and the role for subgroup analysis.

4.4.1 Intention-to-Treat Analysis

Intention-to-treat (ITT) analysis is a technique commonly used in randomized control trials. The definition is as follows: "All patients randomly allocated to one of the treatments in a trial should be analyzed together as representing that treatment, whether or not they completed, or indeed received that treatment" [7]. In other words, ITT compares outcomes between study groups with each participant analyzed according to their randomized group assignment regardless of receiving the assigned treatment, withdrawal from the study, or deviation from the protocol.

An alternative to ITT is "per protocol" analysis, which only evaluates those participants who complied with the assigned treatment. This appears to be an appropriate approach to analysis because participants can only be affected by an intervention they actually received. However, the problem arises when participants who adhere to the study treatment differ from those that are noncompliant or drop out, thus introducing bias [8]. For example, in the Postmenopausal Estrogen/Progestin Intervention (PEPI) Trial, 875 healthy postmenopausal women aged 45–64 years of age who had no known contraindication to hormone therapy were randomly assigned to four different estrogen or estrogen plus progesterone regimens and placebo. Of the 175 women assigned to the unopposed estrogen arm, 41 (23 %) discontinued treatment because of

endometrial hyperplasia, which is a precursor of endometrial cancer [9]. If "per protocol" analysis was performed, these women would have been eliminated from analysis, and the association of estrogen therapy and endometrial cancer may have been missed.

ITT analysis not only minimizes bias but it also maintains the similarities between treatment groups in regard to prognosis. This is the reason for randomization, and this feature may be lost if analysis is not performed on the groups produced by the randomization process. This can be illustrated by the European Coronary Surgery Study Group Trial comparing medical and surgical treatment for stable angina. A total of 768 men under the age of 65 with angina were included in the study (373 men were randomized to medical treatment, and 395 men were randomized to surgical treatment). A total of 26 men assigned to the surgical arm did not undergo surgery, and 50 men assigned to the medical arm underwent surgery. Using ITT analysis, there was no significant difference in mortality between the two groups at 2 years [10]. Alternatively, using "per protocol" analysis, the mortality rate would be 8.4 % for the medical treatment and 4.1 % for the surgical treatment ($P=0.018$) [7]. In "per protocol" analysis, surgery appears to have a falsely low mortality rate.

Despite the advantages of ITT analysis, the major disadvantage is that participants who choose not to take the assigned intervention will be included in the estimate of the effects of that intervention. There is potential for the magnitude of the effect of the treatment to be underestimated if there are a significant number of participants who "cross over" between treatments. For this reason, results of trials are often evaluated using both ITT and "per protocol" analysis, and if both analyses have similar results, the confidence in the trial conclusions is increased. However, if the results from the two analyses are different, the results of ITT analyses dominate because randomization is preserved and bias minimized.

The utilization of ITT analysis has increased over the years. In 1999, Hollis et al. surveyed all reports of randomized controlled trials published in 1997 in the *BMJ*, *Lancet*, *JAMA and New England Journal of Medicine*. A total of 119 (48 %) trials mentioned ITT analysis. Of these, 12 trials excluded any patients who did not start the allocated intervention, and three trials did not analyze all randomized subjects as allocated. The authors concluded that the ITT approach is often inadequately described and inadequately applied, and readers should critically assess the validity of reported ITT analysis [11]. More recently, Gravel et al. conducted a cross-sectional literature review of randomized control trials reported in ten medical journals in 2002. Of the 403 articles, 249 (62 %) reported the use of ITT. Among these, 192 (77 %) clearly analyzed patients according to the groups to which they were randomized. Authors used a modified ITT approach in 23 (9 %) and clearly violated a major component of ITT in 17 (7 %). The approach used in 17 (7 %) was unclear [12].

4.4.2 Subgroup Analysis

Clinical trials are labor intensive and costly, and investigators often use subgroup analysis to extract as much information as possible regarding the effect of a particular

treatment. Subgroup analysis compares subsets of randomized participants. Specifically, investigators compare the treatment effect between one and more subgroups rather than the entire cohort of participants. Subgroups are usually defined based on baseline characteristics. Using subgroup analysis, investigators have the potential to determine in which participants a specific treatment is more (or less) effective (or harmful). For example, a double-blind, placebo-controlled trial was conducted in which the reduction in the incidence of death or hospitalization for cardiovascular reasons with the use of the beta-blocker carvedilol was compared to placebo in patients with heart failure. In subgroup analysis, the investigators further examined whether carvedilol decreased the incidence of cardiovascular events according to the patients' severity of disease, age, sex, left ventricular ejection fraction, 6-min walk, cause of congestive heart failure, systolic blood pressure, and heart rate. Patients treated with carvedilol had a 65 % lower risk of death than those given placebo, and the beneficial effect of carvedilol on survival was consistent in all evaluated subgroups [13].

Even though subgroup analysis allows investigators to identify who, if anyone, benefits from an intervention, care must be utilized in the interpretation of subgroup findings. There are several issues that arise during subgroup analysis [14]:

1. Most trials are not sufficiently powered to detect a difference between treatment groups. Subgroups are by definition smaller that the entire trial cohort. Therefore, if a difference does exist between subgroups, it may not be detected because the size of the trial is not large enough. Investigators may also examine results in a large number of subgroups, thus increasing the likelihood that a difference in treatment effect in a subgroup may be due to chance (type I error).
2. A number of subgroups can be identified based on baseline characteristics, and subgroups can be specified either before or after examination of the data. These two methods are referred to as prespecified subgroup analysis and post hoc analysis, respectively. Prespecified subgroup analysis is planned and documented before any data analysis is performed. Post hoc analysis is often referred to as "data dredging" or "fishing" and can be of particular concern because it can be unclear how many subgroups were analyzed and whether some subgroups were identified secondary to inspection of the data [15].
3. Statistical tests for interaction examine the strength of treatment differences between varying subgroups. This is the best method for making inferences from subgroup analysis. Tests for interaction take into consideration that data available for subgroup analysis is limited. Even though tests for interactions protect investigators from making false or premature claims from subgroup analysis, the test is not routinely used. In a survey of 50 trial reports in 4 major journals conducted by Pocock et al. in 2002, only 15 (43 %) of the 35 reports with subgroup analysis used tests for interaction [14]. A common mistake made by investigators is presenting separate *p*-values for treatment differences within each subgroup. For example, testing the hypothesis that there is no treatment effect in patients younger than 50 years of age and then testing the hypothesis separately in patients older than 50 years of age does not address whether treatment differences vary according to age. Separate subgroup *p*-values can be misleading.

4. The results of subgroup analysis are often overinterpreted by authors and readers, and caution has to be exercised when drawing conclusions. In the same survey conducted by Pocock et al., 21 trials (42 %) claimed to find differences in subgroups that were not compatible with the overall treatment comparison, and 13 of these featured these claims in the summary and/or conclusion [14]. As readers analyze trials that utilize subgroup analysis, biological plausibility, the number of subgroup analyses performed, prespecification of the subgroups, and the trial size have to be considered when drawing conclusions.

Treatment decisions in multiple fields of medicine are directed by the results from randomized clinical trials (RCTs). One field in which there have been hundreds of RCTs is cardiology. Hernandez et al. reviewed 63 cardiovascular RCTs published from 2002 to 2004 in major medical journals. Of the selected RCTs, 39 reported subgroup analysis, and 26 had more than 5 subgroups. Only 14 (35.8 %) prespecified the subgroups, and only 11 (28 %) reported interaction tests. The authors concluded that the reporting of subgroup analysis in cardiovascular RCTs had several shortcomings, including lack of prespecification and testing of a large number of subgroups without the use of tests for interactions. Based on these results, the authors made several recommendations to appropriately perform and interpret subgroup analysis [16]:

1. Specify subgroups in advance with a clear rationale.
2. Use statistical tests for interaction in the full RCT population.
3. Be skeptical if subgroups were not prespecified, not biologically plausible, or no interaction test was performed.
4. Utilize subgroup analysis as a hypothesis-generating tool for future studies.
5. Emphasis should be placed on the overall results, which for the most part are better estimates of treatment effects compared to subgroup effects.

In summary, subgroup analysis is important in clinical trials. The results of subgroup analysis can be used to generate a hypothesis for future studies, but the results must be interpreted with caution, and broad, general conclusion statements should not be made based on subgroup analysis.

4.5 Handling Missing Data

Missing data is a serious problem and has the potential to compromise conclusions drawn from clinical trials. Missing data is defined as “values that are not available and that would be meaningful for analysis if they were observed” [17]. It occurs when participants drop out of a study before its conclusion. Dropout can be secondary to treatment or analysis dropout. Treatment dropout occurs when the assigned treatment is terminated, and analysis dropout occurs when some study measurements are not recorded [18]. If dropout is secondary to the intervention, whether it is treatment dropout or analysis dropout, bias can be introduced into the analysis.

Unfortunately, limited information is available on how to handle missing data. Wood et al. reviewed all randomized trials published between July and December 2001 in the *British Medical Journal*, *Journal of the American Medical Association*, *Lancet*, *and New England Journal of Medicine*. They focused on trial design and how missing outcome data was described and how statistical methods were used to deal with missing data. The conclusion of their review was that missing outcome data is a common problem in randomized controlled trials and it is often inadequately handled in the statistical analysis [19]. To help address this problem, the National Research Council convened the Panel on the Handling of Missing Data in Clinical Trials at the request of the Food and Drug Administration (FDA). The objective of the panel was to prepare "a report with recommendations that would be useful for FDA's development of a guidance for clinical trials on appropriate statistical methods to address missing data for analysis of results" [20]. The recommendations of the panel are summarized below.

The first step in minimizing missing data occurs during the design of the clinical trial. Every effort should be made to clearly define the target population and outcome measures prior to the initiation of the trial. The trial should be designed to maximize adherence to the protocol and ensure participants adhere to follow-up visits and measurements. Little et al. published several suggestions (adopted from the Panel on the Handling of Missing Data in Clinical Trials) for limiting missing data in the design of clinical trials [17, 18]:

1. Target a population that is not adequately served by available treatments and thus have incentive to remain on the study.
2. Include a run-in period in which all participants are initially placed on active treatment. After a specified time, the participants who were adherent to the therapy are randomized to continue active treatment or begin placebo.
3. Allow flexibility in the treatment regimen in order to reduce the dropout rate because of a lack of efficacy or treatment intolerance.
4. Consider add-on designs (study treatment is added to an existing treatment).
5. Shorter follow-up periods for the primary outcome.
6. Allow the use of rescue medications.
7. Consider a randomized withdrawal design to assess long-term efficacy (participants who have received study treatment without dropping out are randomized to continue to receive the treatment or switch to placebo).
8. Try to avoid using outcome measures that are likely to lead to substantial missing data.

Another important factor to take into consideration during the design phase is how missing data will affect the power of the trial. Most investigators "inflate" the initial sample size to account for anticipated missing data.

Even when investigators take every step to minimize missing data during the design of the trial, every participant will not follow their assigned intervention to the completion of the trial. The question then presents itself in regard to which data, if any, should be collected for participants who do not complete the assigned treatment. Investigators have two opposing views. Some believe that participants who

do not complete the assigned treatment are no longer relevant to the study. The opposing view is that continued data collection may be informative and can potentially allow for the ability to analyze end points for all participants and explore whether assigned therapy effects the use and efficacy of subsequent therapies [18].

After taking steps to minimize missing data during the design of the clinical trial, attention must then be turned to minimizing missing data during the conduct of the trial. Little et al. once again have several suggestions (also adopted from the Panel on the Handling of Missing Data in Clinical Trials) for limiting missing data during conduct of the trial [17, 18]:

1. Select investigators who have good track records enrolling participants, following participants, and collecting complete data.
2. Set acceptable rates for missing data in the study protocol.
3. Provide incentives (as long as they comply with ethical requirements) to investigators and participants for completeness of data collection.
4. Minimize the participant inconvenience and burden associated with data collection.
5. Provide effective treatment to participants after the trial.
6. Train investigators and their research staff on the negative impact of missing data.
7. Train investigators and their research staff on the informed consent process as a tool for encouraging complete data.
8. Monitor missing data during the trial.

Unfortunately, there is no universal method for handling missing data during data analysis. The Panel on the Handling of Missing Data in Clinical Trials identified four different methods to adjust for missing data: complete-case analysis, single imputation methods, estimating equation methods, and methods based on a statistical method. Complete-case analysis excludes participants with missing data from the analysis. Single imputation methods fill in a value for each missing value using methods such as the last observation or baseline observation carried forward. Estimating equation methods weigh complete cases by the inverse of an estimate of the probability of being observed, and methods based on statistical methods include maximum likelihood, Bayesian methods, and multiple imputations. In general, the panel favored estimating equation methods and methods based on a statistical model for the data [17].

In summary, missing data is a major issue that has to be addressed in the design and analysis of clinical trials. Missing data can lead to bias and affect the interpretation of trial results. Therefore, it is important to try to minimize missing data during the design and conduct of clinical trials.

4.6 CONSORT Statement

In an effort to facilitate the interpretation of data from randomized trial and to facilitate their complete and transparent reporting such that some of the issues described above can be deliberated during the interpretation of the results, scientists and

editors have developed the Consolidated Standards of Reporting Trials (CONSORT) statement [21]. It is comprised of a 25-item checklist and a flow diagram focusing on reporting how the trial was designed, analyzed, and interpreted. While use of the CONSORT statement improves the communication of the study design and its findings, it is important to understand these key issues for the interpretation of data.

4.7 Conclusion

RCTs provide the foundation for evidence-based medicine. The design and implementation of RCTs are labor extensive and expensive; therefore, it is important that investigators have a clear understanding of the design and implementation of clinical trials and the analysis of the results obtained during a clinical trial. This chapter provided a brief review of common statistical methods utilized to express the results of clinical trials and common issues associated with data analysis.

References

1. Friedman LM, Furberg C, DeMets DL. Fundamentals of clinical trials. 3rd ed. New York: Springer; 1998. xviii, 361 p.
2. Dawson B, Trapp RG. Basic & clinical biostatistics. 4th ed. New York: Lange Medical Books-McGraw-Hill, Medical Pub. Division; 2004. x, 438 p.
3. Wang D, Bakhai A, Maffulli N. A primer for statistical analysis of clinical trials. Arthroscopy. 2003;19(8):874–81.
4. Dawson B, Trapp RG. Basic & clinical biostatistics. 4th ed. New York: Lange Medical Books/McGraw-Hill; 2004.
5. Straus SE. Evidence-based medicine : how to practice and teach it. 4th ed. Edinburgh: Elsevier Churchill Livingstone; 2011. xvii, 293 p.
6. Troidl H. Principles and practice of research: strategies for surgical investigators. Berlin/New York: Springer; 1986. xvii, 380 p.
7. Newell DJ. Intention-to-treat analysis: implications for quantitative and qualitative research. Int J Epidemiol. 1992;21(5):837–41.
8. Hulley SB. Designing clinical research: an epidemiologic approach. 2nd ed. Philadelphia: Lippincott Williams & Wilkins; 2001. xv, 336 p.
9. Effects of estrogen or estrogen/progestin regimens on heart disease risk factors in postmenopausal women. The Postmenopausal Estrogen/Progestin Interventions (PEPI) Trial. The Writing Group for the PEPI Trial. JAMA. 1995;273(3):199–208.
10. Coronary-artery bypass surgery in stable angina pectoris: survival at two years. European Coronary Surgery Study Group. Lancet. 1979;1(8122):889–93.
11. Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. BMJ. 1999;319(7211):670–4.
12. Gravel J, Opatrny L, Shapiro S. The intention-to-treat approach in randomized controlled trials: are authors saying what they do and doing what they say? Clin Trials. 2007;4(4):350–6.
13. Packer M, et al. The effect of carvedilol on morbidity and mortality in patients with chronic heart failure. U.S. Carvedilol Heart Failure Study Group. N Engl J Med. 1996;334(21):1349–55.
14. Pocock SJ, et al. Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems. Stat Med. 2002;21(19):2917–30.

15. Wang R, et al. Statistics in medicine–reporting of subgroup analyses in clinical trials. N Engl J Med. 2007;357(21):2189–94.
16. Hernandez AV, et al. Subgroup analyses in therapeutic cardiovascular clinical trials: are most of them misleading? Am Heart J. 2006;151(2):257–64.
17. Little RJ, et al. The prevention and treatment of missing data in clinical trials. N Engl J Med. 2012;367(14):1355–60.
18. Little RJ, et al. The design and conduct of clinical trials to limit missing data. Stat Med. 2012;31(28):3433–43.
19. Wood AM, White IR, Thompson SG. Are missing outcome data adequately handled? A review of published randomized controlled trials in major medical journals. Clin Trials. 2004;1(4):368–76.
20. Council NR. The prevention and treatment of missing data in clinical trials. Washington, DC: National Academic Press; 2010.
21. Moher D, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. J Clin Epidemiol. 2010;63(8):e1–37.

Chapter 5
Data Safety Monitoring Boards

Kamran Idrees, Fei Ye, Yu Shyr, and Nipun Merchant

5.1 Introduction

Randomized control trials (RCTs) provide essential evidence on the efficacy and safety of medical interventions. Monitoring of safety outcomes is always needed during any trial involving human participants that entails more than minimal risk. The nature and degree of this monitoring is important for several reasons. Firstly, it is critical to safeguard the interests of the study participants. Secondly, it ensures the scientific validity and integrity of the study. Lastly, it allows for early termination of the study based on interim results demonstrating futility or positive efficacy to study participants. In many cases, monitoring can be adequately performed by investigators and sponsors [1]. However, when this role is performed by a formal independent multidisciplinary group of specialists, it is referred to as Data and Safety Monitoring Board (DSMB).

A DSMB [also known as data monitoring committees (DMCs) or Data and Safety Monitoring Committees (DSMCs)] was first established by the National Institutes of

K. Idrees, MD • N. Merchant, MD (✉)
Division of Surgical Oncology and Endocrine Surgery, Department of Surgery,
Vanderbilt University, 2220 Pierce Avenue, 597 Preston Research Building,
Nashville, TN 37232-6860, USA
e-mail: kamran.idrees@vanderbilt.edu; nipun.merchant@vanderbilt.edu

F. Ye, PhD
Division of Cancer Biostatistics, Department of Biostatistics,
Vanderbilt Center for Quantitative Sciences, 495 Preston Research Building,
Nashville, TN 37232-6848, USA
e-mail: fei.ye@vanderbilt.edu

Y. Shyr, PhD
Department of Biostatistics, Vanderbilt Center for Quantitative Sciences,
571 Preston Research Building, Nashville, TN 37232-6848, USA
e-mail: yu.shyr@vanderbilt.edu

T.M. Pawlik, J.A. Sosa (eds.), *Success in Academic Surgery: Clinical Trials*,
Success in Academic Surgery, DOI 10.1007/978-1-4471-4679-7_5,

Health (NIH) in the 1960s as an external task force report known as the Greenberg report [2]. Initially used in trials funded by government agencies in which the goals of the trial were to reduce mortality and/or decrease major morbidity, the DSMB's role was to carefully monitor the accumulating results, to consider whether modifications of trial conduct were needed, and to make recommendations to the investigators regarding continuation or termination of the trial [2, 3]. With the increase in industry sponsored trials by pharmaceutical and medical device companies as well as increased public awareness regarding potential bias, DSMBs are frequently incorporated in the conduct of trials. DSMB members consist of a group of experts who are independent of the trial, without any conflict of interest, and who can make recommendations regarding the conduct of the trial on the basis of emerging data while minimizing unwanted influences on their judgment [4].

5.2 When Is a DSMB Required?

All clinical trials necessitate safety monitoring, but not all randomized controlled trials require monitoring by an independent board or DSMB. For most phase I and II clinical trials, as long as safety of the participants and trial integrity is assured by investigators and study sponsors, a DSMB is not required. These trials are usually conducted at a single institution, are small and of short duration; therefore, establishing a DSMB may not be practical. However, if a novel drug, therapy, or device with high or unknown risk profile is used in a phase I trial, monitoring by the principal investigator (PI) or institutional review board (IRB) might be insufficient, and a DSMB may be required. Similarly, in a phase II randomized trial, if the assignment to the control or intervention group is blinded, then the investigator would not be able to perform the monitoring functions. Another person or group would be required to do the monitoring in such a trial to assure that investigators remain blinded.

DSMBs are generally established for late-phase clinical trials (III and IV) which are usually large, randomized, double-blinded, and conducted in multiple institutions. There are several reasons for having a DSMB for late phase. Firstly, these large trials are of long duration and therefore increase risks and safety concerns because of greater and prolonged exposure which may cause adverse events that might not be initially recognized. Secondly, these trials are usually double-blinded, which makes it impossible for the investigator to perform monitoring. Thirdly, many late-phase trials are conducted in multiple centers making monitoring by a single investigator difficult. Lastly, and most importantly, the purpose of a late-phase clinical trial is to provide an answer with the potential to profoundly alter clinical practice. Employing a DSMB provides scientific merit to the study and allows for a decision to be made by someone who is not involved (directly or indirectly) in the research, thereby minimizing the risk of continuing the study longer than appropriate or terminating it sooner than reasonable. It also avoids biases (conscious or subconscious) on the part of the investigator to see a particular result [5, 6]. General criteria used to determine whether a DSMB is required for a particular trial is shown in Table 5.1.

Table 5.1 Criteria that necessitate a DSMB

1. High risk or vulnerable populations
2. Trials investigating new interventions with few or no safety data available
3. Trials with primary endpoints such as mortality or major morbidity
4. Randomized multicenter or international trials
5. Planned interim analyses with the potential for early termination
6. Trials with a large sample size

5.3 Organizing a DSMB

A DSMB typically consists of one or more clinicians knowledgeable in the field of investigation as well as biostatisticians or clinical trial methodologists [4]. The expertise required for each DSMB varies according to the design and complexity of the trial. The clinicians could be physicians, nurses or other allied health specialists with expertise in the particular disease under study. They are responsible for reviewing the adverse event monitoring plans and serious adverse events. These experts also determine the appropriateness of measurements used for data reporting and can advise on quality control and/or assurance associated with laboratory or imaging measurements and instrumentation. Biostatisticians with experience in clinical trial design or clinical trial methodologists should be included in DSMB composition and tasked with review of the analysis plan and sequential analysis of interim trial data. Other specialists that may be valuable for some trials include clinical pharmacologists, bioethicists, and patient advocacy experts. Bioethicists on DSMBs provide expertise on informed consent and to ensure protection of vulnerable populations. A consumer or community advocate (often a current or former patient, a parent, or family member of someone with the disease) may also provide a helpful perspective [7].

Although there is no set number of members required for a DSMB, a general principle is that it should be as small as possible while encompassing all relevant expertise (odd numbers are preferred to avoid tie votes) [8]. One of the individuals is selected to serve as DSMB chairperson; he/she is the point person for the DSMB and is responsible for overseeing the meetings. Although there is no official requirement for the selection of the DSMB chairman, it is desirable to have someone with previous DSMB experience who is an accomplished group leader with good interpersonal and organizational skills.

DSMB members are usually appointed either by principal investigators or trial sponsors, depending on the type of trial, the phase of trial, and sponsoring agency. For early-phase NIH-sponsored trials, the principal investigator typically appoints the members of the DSMB. For large, multicenter phase III and IV trials, the funding institute often nominates the board. For industry-sponsored trials, the pattern varies from either the company appointing the members or it adopts a “hands-off” approach and lets an external group do it. The members of DSMB should be independent of the sponsor of the study and manufacturer of the product evaluated. They should also not

be related to the study's investigative group. The DSMB members should have no financial, scientific, or any other conflict of interest with the trial. Absence of major conflict of interest is an essential requirement for all DSMB members to ensure the independence of the board. Financial conflicts (e.g., substantial stock holdings or yearly stipend from a sponsor) have received the most attention in the media, but other conflicts (e.g., strong intellectual investment or personal relationships) can also be important. For example, an individual whose initial concept or early work is being tested in the study could gain tremendously if the original research concept is validated and therefore may be hesitant to consider early termination of a trial in case of severe adverse events, compared to an individual who is completely independent. Similarly, the members of DSMB should not receive coauthorship on articles or promotions based on trial results because it creates conflict of interest.

5.4 Responsibilities of DSMB

The overarching purpose of a DSMB is to protect the safety of study participants and to give credibility and validity to the results of the study. The fundamental responsibility of every DSMB is to make recommendations to the trial's sponsor and/or steering committee concerning the appropriateness of trial continuation. The DSMB has broad responsibilities which are summarized in Table 5.2.

Based on the accumulating data and interim analysis, DSMBs typically make one of the following recommendations: continuation of the trial as planned; continuation of trial with modifications; early termination of the trial for either futility or adverse events; early termination of the trial based on unequivocal efficacy; or extension of the trial.

Commonly, DSMB is asked to review and approve the study protocol before initiation. Although DSMB is not responsible for study approval (which is done by an IRB), the purpose of this practice is to ensure that the board members are in agreement with the study investigators and sponsors about the acceptability of study design, mainly the statistical monitoring plan and any criteria for early termination.

Table 5.2 Responsibilities of a data safety monitoring board

1. Review research protocol, informed consent documents, and plans for data safety and monitoring
2. Give priority to monitoring harm or adverse effects on study participants
3. Evaluate performance of individual centers including periodic assessments of data quality, participants recruitment, accrual, and retention
4. Analyze interim results and evidence of efficacy or adverse events
5. Consider any newly available external evidence such as scientific or therapeutic developments that can impact the ethics of the trial or safety of trial participants
6. Maintain confidentiality of the trial data
7. Report on the safety and progress of the trial
8. Make recommendations to the sponsor and investigators regarding continuation, modification, or
9. Assist in resolution of the problems reported by investigators

It is a delicate balance between protecting study participants and ensuring the reliability of the trial results when selecting criteria for early termination. For example, in a trial evaluating a potentially lifesaving treatment of a seriously ill patient, a less stringent criterion might be used to allow quick availability of a superior treatment. On the other hand, in evaluating a new vaccine, it is important to collect data over a longer duration to ensure that the vaccine can be safely administered to millions of healthy individuals and in such a scenario early termination is extremely rare. Occasionally, sponsors or investigators may ask DSMB to release interim data because of safety concerns arising in other investigational treatment trials or requests by regulatory authorities. Since DSMB protects the confidentiality of the interim data, it determines on an individual basis whether to grant such a request while safeguarding the integrity of the trial.

5.5 Workings of a DSMB

The DSMB functions with written operating guidelines which are typically prepared by the sponsor and approved by the DSMB during the first meeting [4, 9]. This document outlines the conduct and frequency of meetings, meeting formats, documentation and reporting procedures of all meetings, discussions and decisions, and ensures that IRBs are promptly informed of meeting recommendations. The number of interim analyses should be adequately planned and should be conducted on predefined outcomes. DSMB meetings can be divided into "open" or "closed" sessions, depending on the phase of the study. During the open session, study progress is reviewed, and therefore it can include study investigators, sponsors, monitors, and regulators who can answer questions raised by the DSMB related to the trial's conduct. During the closed session, study efficacy and safety data is reviewed and typically includes only DSMB members. The objective is to avoid unblinding the investigators or industry sponsors and to prevent them from developing biases.

The DSMB receives two kinds of reports to evaluate, one regarding process and the other relating to study outcomes. The study outcome report consists of primary and secondary outcome variables as well as adverse events. This outcome report is provided to the DSMB members only. The process reports usually consist of study participant recruitment, accrual and retention status, data quality and timeliness and performance of trial sites. These process reports, including DSMB recommendations, are shared with trial investigators, sponsors, and regulatory authorities.

5.6 Statistical Issues Related to DSMBs

The DSMB must review accumulating data periodically in order to assess whether an important safety issue has arisen (e.g., severe side effects), or whether the intervention under study is showing a substantial beneficial effect earlier than expected. The required frequency of these reviews depends on the disease and the specific

interventions. Most DSMBs hold meetings two to four times a year. While these interim reviews are conducted, the repeated statistical evaluation of data must be performed with caution, especially in the early stages of a trial when the number of participants and the number of events are relatively small. In addition to clinicians with expertise in relevant clinical specialties, most DSMBs have at least one biostatistician who is knowledgeable about statistical methods for clinical trials and sequential analysis of trial data.

5.6.1 Interim Analysis

An interim analysis is planned that intends to compare treatment arms at any time prior to formal completion of a trial. When interim analyses are necessary, they should be preplanned for early stopping (either for futility or for positive efficacy), or in the case of adaptive designs where a possible modification of the study design is based on unblinded interim data. Interim data, whether blinded or unblinded, should be analyzed by a statistical group that is independent of the sponsor and investigators—that is, the group is not otherwise involved in the trial design or conduct and has no financial connections to the sponsor or other trial organizers.

5.6.2 Primary versus Secondary Endpoints

A clinical trial endpoint is defined as a measure of efficacy outcome used to decide whether the null hypothesis of a clinical trial should be accepted or rejected. In a clinical trial, the null hypothesis states that there is no statistically significant difference between two treatments or strategies being compared with respect to the endpoint measure chosen. To avoid multiplicity concerns, most trials designate a primary endpoint. In some cases, a composite primary endpoint, which represents the occurrence of at least one of several different important outcomes, is used to overcome the difficulties associated with designing a study around a single event of interest (e.g., death). The size of a trial is determined by the power needed to detect a difference in this primary endpoint. The primary endpoint is important for the study design because the sample size required in order to adequately power a study is dependent upon the number of primary events that are expected to occur over a given time period. A DSMB should maintain focus on the primary endpoint, particularly for purposes of early termination, because it is often difficult to make a decision by considering endpoints other than the one designated as primary.

Secondary endpoints ask other relevant questions about the same study, for example, whether there is also a reduction in disease recurrence or whether the new treatment improves the overall quality of life of the treated patients. When

secondary endpoints are also important, the trial must be powered sufficiently to detect a difference in both primary and secondary endpoints.

5.6.3 Control of Type I Errors

At each interim stage, all the accumulated data up to that point are analyzed, and a decision is made about whether the trial will be stopped or continued. A major concern when emerging data is repeatedly analyzed is that the type I error (false positive) rate may be inflated if adjustment is not made for multiple looks at the data (frequentist approaches). Typically, the monitoring plan will specify a statistical approach that permits multiple interim analyses while maintaining the type I error rate at the desired level. These approaches usually generate boundaries for beneficial interim estimates that indicate the magnitude of benefit needed to support stopping the trial at interim points prior to its planned completion, while maintaining the desired overall probability of type I errors. Such boundaries can serve as useful guidelines for the DSMB in making recommendations regarding continued accrual to and conduct of the trial. If the implementation of a group sequential trial involves unblinding the interim results and analyzing the interim treatment effect, it can raise concerns of potential bias. The FDA (Food and Drug Administration) recommends that the analyses be carried out either external to the trial's sponsor or by a group within the sponsor that is unequivocally separated from all other parties to the trial. In addition, some group sequential design approaches may not be able to control the overall type I error rate if the target patient population has been shifted due to additional adaptations or protocol amendments.

5.6.4 Group Sequential Designs

Group sequential designs are commonly applied to facilitate the conduct of interim analysis. As previously mentioned, repeatedly examining accumulating data increases the chance of falsely claiming a beneficial effect if we use the standard critical value (such as 1.96 for a two-sided 0.05-level test) for a test of significance. Group sequential design offers opportunities for early termination of a trial for either safety or efficacy reasons, allowing the sample size to be reduced. At each interim stage, all the accumulated data up to that point are analyzed, and a decision is made about whether the trial will be stopped or continued. Such decisions are made at interim analyses, and various stopping rules are available in the literature. Most frequentist methods attempt to control the overall type I error rate and to terminate the trial when there is neither enough beneficial treatment effect nor sufficient efficacy observed, using approaches such as alpha-spending rules. Because the number, methods, and consequences of these analyses affect the interpretation of the trial, all interim analyses should be carefully planned in advance and described

in the protocol. Statistical methods should be fully specified in advance of the availability of data on treatment assignments and treatment outcomes.

5.6.5 Sample Size

The number of participants in a clinical trial should always be large enough to provide a reliable answer to the questions addressed. This number is usually determined by the primary endpoint of the trial. The method by which the sample size is calculated should be clearly specified in the protocol; the basis of these estimates should also be given. Using a frequentist approach in determining the appropriate sample size, the following items should be specified: a primary endpoint, the test statistic, the null hypothesis, the alternative hypothesis at the chosen dosage(s), and the type I and type II error rate constraints, as well as the approach to dealing with patient withdrawals and protocol violations. All else being equal, statistical power increases with the increasing number of participants enrolled in the trial. However, the effort to increase power must be balanced with the greater costs associated with larger trials.

It is important to investigate the sensitivity of the sample size calculation to a variety of deviations from the assumptions being made. The sample size of an equivalence trial or a non-inferiority trial is usually based on the confidence interval for the treatment difference under the alternative hypothesis. The exact sample size in a group sequential trial cannot be fixed in advance because it depends upon the chosen stopping guideline and the true treatment difference. However, the distribution of the sample size should be provided, usually embodied in the expected and maximum sample sizes. When event rates are lower than anticipated or variability is larger than expected, methods for sample size reestimation can be used (see Sect. 5.6.8).

5.6.6 Efficacy and Futility

Interim analyses are performed for safety, efficacy (the new treatment performs overwhelmingly better than the control), or futility (the new treatment is unlikely to be better than the control) reasons. For trials that may be terminated because of safety concerns, timely communication with the FDA is often required. The investigators should initiate discussion with the FDA prior to early termination of any trial implemented specifically to investigate a potential safety concern. For trials that may be terminated early because a substantial benefit has been observed (efficacy), however, consideration should still be given to the adequacy of data with regard to other issues such as safety, duration of benefit, outcomes in important subgroups, and important secondary endpoints. When statistical analyses suggest that the treatment is unlikely to meet the objectives of the trial based on the interim data, a DSMB may recommend early termination of the trial on the grounds that

there is no basis for continuing enrollment and/or follow-up. In this case, the type II error should be examined for estimating the chance of making a false negative conclusion. Nevertheless, protection of a type I error is important even when the decision of early termination is made only for futility reasons, since interim review of outcome data always raises the possibility of a type I error.

5.6.7 Analysis Methods: Statistical Philosophies

Three major philosophies of statistical methods coexist in clinical trial design and analysis today: frequentist, Bayesian, and likelihood approaches. The distinction between frequentist and Bayesian approaches is partly the result of viewing parameters to be estimated as random variables (Bayesian) or fixed constant of nature (frequentist). In a typical phase II trial design, the frequentist approaches optimize some criterion subject to error constraints at null and alternative values of the response rate. The Bayesian approaches incorporate prior information updated by data as they accumulate and the design calls for termination at any interim analysis when an observed persuasion probability exceeds its critical value. The so-called persuasion probability is based on the Bayesian posterior probability that the experimental treatment is superior to the standard. To assess frequentist characteristics such as type I and type II error rates of a Bayesian design, an intensive simulation study that considers a wide range of possible scenarios should be carried out. The main purpose is to evaluate how robust and reliable the Bayesian design is under different circumstances, especially when compared with a standard frequentist design. The likelihood approaches use the likelihood principal and the statistical likelihood function to make inference. The likelihood has a natural interpretation as a quantification of relative evidence. However, for many reasons it is not as widely used in clinical trial design as the frequentist or Bayesian approaches.

5.6.8 Adaptive Design

Adaptive trial designs allow for a prospectively planned modification of certain aspect(s) of the trial design at interim stages, aiming to improve the efficacy of a trial and increase the chance of success, while enhancing investigators' understanding of the effect of the treatment. Types of design adaptations include modifications of study eligibility criteria, randomization procedure, sample size, primary and secondary endpoints, treatment allocation, number of interim analyses, dosage levels, and duration, as well as methods of statistical analysis.

Many critical parameters used in planning a trial are estimated based on certain assumptions, such as response rates, standard deviations, and population means, because of an incomplete or inadequate understanding of these elements. Consequently, a trial may fail to achieve its goal when these prespecified estimates

or assumptions substantially deviate from the truth. In many cases, investigators have a good understanding of an investigational drug only after the data have been collected and unblinded. It is not uncommon for investigators to discover only then that the dosage was subtherapeutic or that some arms of the trial proved unnecessary and could have been dropped early in the trial. However, within the structure of a nonadaptive clinical trial design, investigators can do little to address such limitations without harming the validity of the trial.

Over the past 20 years, such concerns have stimulated a tremendous effort to improve the efficiency of trial designs. A common theme has been a move toward adaptive designs. Use of the term "adaptive" has a long history in clinical trial literature. Cornfield and colleagues proposed an analytic approach for adaptive trials motivated by the two-armed bandit model, which aims to maximize the number of patients assigned to the more promising of two treatments [10]. Zelen first introduced the concept of the "play-the-winner" rule for the same purpose [11]. Wei and Durham proposed their play-the-winner rule for the randomization procedure in a sequential trial setting as an improvement of Zelen's design [12]. Today, the umbrella of adaptive design covers many approaches to introducing some degree of flexibility into a trial, ranging from the basic 3 + 3 phase 1 trial design for dosage finding to cutting-edge biomarker adaptive design. Some researchers categorize adaptive trial designs based on the rules for adaptations. There are roughly four categories: allocation rules, sampling rules, stopping rules, and decision rules. Allocation rules define how patients will be allocated to different arms in a trial. The sampling rule defines how many patients will be enrolled at the next stage. The stopping rule considers when to stop the trial for reasons such as efficacy, futility, harm, or safety. The decision rule refers to design modifications that are not covered by the previous three rules, including change of endpoint, trial hypothesis, and statistical analysis plan. The most widely used adaptive design methods include adaptive dosage-finding design, adaptive randomization design, sample size re-estimation design, group sequential design, adaptive seamless phase 2/3 design, adaptive treatment selection design, and biomarker design. The common goal is to incorporate learning from running a trial into its later stages, so that the trial design can be corrected or improved when evidence suggests that misspecifications have occurred in the trial's planning phase. Appropriately, the FDA draft guidance on adaptive clinical trials defines an adaptive design clinical study as "a study that includes a prospectively planned opportunity for modification of one or more specified aspects of the study design and hypotheses based on analysis of data (usually interim data) from subjects in the study."

Adaptive design offers the potential to yield more information about the investigational treatment than would otherwise be feasible within the time and resources allowed for a particular trial. An adaptive design may allow investigators to discontinue the data collection of one or more arms when the current accumulated data are shown to be ineffective for the arm(s), thereby reducing costs and time spent on treatments that are not promising, based on the "learn-and-confirm" model, without decreasing the useful information gained from the overall trial. In addition, adaptive design trials may improve the understanding of the dose–response relationship using approaches such as continual reassessment methods [13, 14].

5.6.9 Reporting on DSMB Activities

A description of the DSMB and its activities, such as names of DSMB members with their respective affiliations and areas of expertise, DSMB roles, interim analyses (frequency, planned or *ad hoc*, statistical methods used), and termination guidelines defined *a priori*, should be detailed in manuscripts. This clarifies the study monitoring process as well as allows readers the DSMBs' impact on the scientific merit of the trial results.

5.7 Summary

All clinical trials that involve risk to study participants require a data safety monitoring plan. The extent and approach varies depending on the nature of the clinical trials and the type of intervention. Many late-phase clinical trials need a formal independent monitoring group such as a DSMB, whose primary task is to protect trial participants and lend credibility to the trial results. DSMB membership should be broad enough to include clinical and methodological expertise, whose work should be guided by a detailed charter that benefits all aspects of a clinical trial. This peer review process allows safety of trial subjects and produces the highest quality data providing vital results that can profoundly alter the practice of medicine or public health policy. There is no substitute for an experienced, independent, unbiased data safety monitoring board dealing with the complex issues involved in making the decision to terminate or modify a clinical trial.

References

1. Friedman LMFC, DeMets DL. Fundamentals of clinical trials. 3rd ed. New York: Springer; 1998.
2. Organization, review, and administration of cooperative studies (Greenberg Report): a report from the Heart Special Project Committee to the National Advisory Heart Council, May 1967. Control Clin Trials. 1988;9(2):137–48.
3. Practical issues in data monitoring of clinical trials'. Proceedings of a workshop. Bethesda, Maryland, 27–28 January 1992. Stat Med. 1993;12(5–6):415–616.
4. Ellenberg SFT, DeMets D. Data monitoring committees in clinical trials: a practical perspective. 1st ed. West Sussex: Wiley; 2002.
5. Wittes J. Behind closed doors: the data monitoring board in randomized clinical trials. Stat Med. 1993;12(5–6):419–24.
6. Shaw LW, Chalmers TC. Ethics in cooperative clinical trials. Ann N Y Acad Sci. 1970;169(2):487–95.
7. Smith MA, Ungerleider RS, Korn EL, Rubinstein L, Simon R. Role of independent data-monitoring committees in randomized clinical trials sponsored by the National Cancer Institute. J Clin Oncol. 1997;15(7):2736–43.
8. Cairns JA, Hallstrom A, Held P. Should all trials have a data safety and monitoring committee? Am Heart J. 2001;141(1):156–63.

9. Damocles Study Group NHSHTAP. A proposed charter for clinical trial data monitoring committees: helping them to do their job well. Lancet. 2005;365(9460):711–22.
10. Cornfield J, Halperin M, Greenhou SW. An adaptive procedure for sequential clinical trials. J Am Stat Assoc. 1969;64(327):759–70.
11. Zelen M. Play winner rule and controlled clinical trial. Am Stat Assoc J. 1969;64(325): 131–46.
12. Wei LJ, Durham S. Randomized play-winner rule in medical trials. J Am Stat Assoc. 1978;73(364):840–3.
13. O'Quigley J, Pepe M, Fisher L. Continual reassessment method: a practical design for phase 1 clinical trials in cancer. Biometrics. 1990;46(1):33–48.
14. O'Quigley J, Shen LZ. Continual reassessment method: a likelihood approach. Biometrics. 1996;52(2):673–84.

Chapter 6
Planning for Data Monitoring and Audits

Lisa Jacobs

Auditing and monitoring of clinical trials is crucial to ensure the integrity of the research process. An auditing program will evaluate a study for compliance with federal regulatory requirements, accountability of research drugs, and review of submitted patient data. Each of these components is evaluated according to specific federal guidelines for auditing. The most successful audits occur when preparation for the audit starts at the trial design phase and the principal investigator conducts the trial with oversight of the research process and staff. Monitoring is generally completed for pharmaceutical sponsored trials and includes the same components as auditing. For multicenter studies, the study sponsor will establish the quality assurance program to be used. For single institution studies, the need for auditing or monitoring is established by the institutional research policies. Larger research institutions will have a group that performs audits of the institutional protocols that generally functions through the Institutional Review Board (IRB).

6.1 What Is the Difference Between Monitoring and Auditing?

The quality assurance program selected for a research study may be completed either through a monitoring process or an auditing process. Monitoring and auditing are different processes to achieve the common goal of assuring the safety of patients enrolled in clinical trials and the integrity of the research process and data collected. Both methods of quality assurance evaluate compliance

L. Jacobs, MD
Department of Surgery, Johns Hopkins University,
600 N. Wolfe Street, Carnegie 681, Baltimore, MA 21287, USA
e-mail: ljacob14@jhmi.edu

T.M. Pawlik, J.A. Sosa (eds.), *Success in Academic Surgery: Clinical Trials*,
Success in Academic Surgery, DOI 10.1007/978-1-4471-4679-7_6,

with regulatory requirements, IRB functions, informed consent content and process, compliance with study policies and procedures, participant eligibility, control of research drugs and devices, and the accuracy of data submitted [1]. All studies completed under an IND or IDE must undergo monitoring as required by the FDA [2].

The purpose of monitoring is to document the progress of the clinical trial from all aspects, including regulatory, data submission, adherence to study protocol and procedures, and includes a data safety monitoring plan. Monitoring is completed by the study sponsor and requires 100 % verification of study data. This is accomplished through regularly scheduled periodic visits to the sight to oversee the study. Depending on study accrual and the frequency of data collection, these visits may occur as frequently as monthly.

Auditing is an intermittent evaluation of study progress that evaluates a subset of the data collected at a specific time point in the study. Many studies require that 10 % of the patients accrued are audited. It provides a determination of compliance with study policies and procedures, regulatory requirements, and confirms crucial data points through source verification. Audits are generally conducted within the first year of the first patient enrollment and then every 3 years if the site has a successful audit. However, this may be modified at the discretion of the sponsor depending on site performance and accrual.

6.2 Role of the IRB

Guidelines from the FDA state that the IRB review is completed "to assure, both in advance and by periodic review, that appropriate steps are taken to protect the rights and welfare of humans participating as subjects in the research. To accomplish this purpose, IRBs use a group process to review research protocols and related materials (e.g., informed consent documents and investigator brochures) to ensure protection of the rights and welfare of human subjects of research" [3]. The IRB evaluates research protocols for safety and clarity. Some institutions use the IRB to evaluate studies for scientific merit and other institutions use other research committees to serve this function. A review of scientific merit will address ability of the study design to achieve the primary and secondary objectives and feasibility of the study. The primary and secondary objectives and the inclusion and exclusion criteria must be clearly stated. The study calendar identifies the timing of the study procedures and the study reports that are to be submitted. In designing a clinical trial, clear delineation of the components to the study are necessary to establish an effective monitoring or auditing program. The quality assurance program will be designed using the study parameters listed above to ensure research compliance and to ensure that the integrity of the study will be maintained allowing the primary and secondary objectives to be met. The audit or monitoring team works with the IRB representatives, the research staff, and the principal investigator to provide an assessment of the research integrity.

6.3 How Study Design Plays a Role in Quality Assurance

All quality assurance programs start with the study document and then create an auditing or monitoring program from that document that is in compliance with federal guidelines. The design of the study document plays a role in the future success of the quality assurance program. The integrity of the study is affected by the adherence to study protocols and procedures by the research sites. If a study is written in an ambiguous manner that is open to interpretation by the sites, there will not be consistency in the completion of the study parameters. The auditors or monitors will evaluate the research site's compliance with study procedures using their understanding of the requirements of those ambiguous parameters. For example, a study may require two different imaging modalities to assess response to treatment such as CT scan and plain x-ray. If the study does not specifically define which imaging study is to be used to determine response, some sites will use the plain x-ray and others will use the CT scan. This will lead to variability in the study results, and if the auditors have been instructed that the response criteria are based on CT scan, then those sites that used plain x-ray will receive a deficiency. Similar issues arise when inclusion criteria and exclusion criteria are not well defined. For example, many studies require documentation that female participants are not pregnant. The study must specify what documentation is required for women who are postmenopausal or perimenopausal. If this is not adequately clarified, sites would have to complete pregnancy tests on elderly women to document definitively that the participant is not pregnant. All components of a study must be carefully evaluated from a quality assurance perspective for clarity to ensure that the proper participants are recruited and that the desired data points are collected in the manner intended by the study plan. This will enhance the integrity of the study and improve the quality assurance performance of all of the sites.

6.4 Selection of Trials to Participate In

An investigator will be most successful in completion of a study when he has an interest in the topic being investigated. It is also crucial that the investigator have equipoise on the study question. If the investigator already believes that the study question has been answered, he is less likely to enroll subjects on the study particularly if there is a possibility that his patient will not receive a treatment that he feels is already established as more beneficial. Studies are also more successful when the investigators were involved in the design of the study. Participation in the design allows the investigator to participate in the selection of the endpoints and the variables to be tested.

When an investigator is evaluating a study for possible participation, it is important to not only evaluate the study for scientific merit but also for feasibility and clarity. If components of the study are unclear, the investigator should clarify those

with the study sponsor before initiating the study. As the study moves forward, the investigator should continue to communicate with the study sponsor if there are any parameters that are ambiguous or unclear to the investigator. The experience of the organization sponsoring the study is also important to consider. Research organizations with less experience will write protocols that are not clear and therefore require more amendments. Each amendment will have to be submitted and reviewed by the IRB which increases the burden on the research site both for workload and ensuring compliance with the protocol changes. Completion of a clinical trial is a joint effort between the site and the study sponsors. The personnel at each group also play a role in successful completion of the study and in the overall data quality. The study personnel at the site may include the study coordinator, data manager, IRB staff, and the principal investigator. It is the responsibility of the principal investigator to ensure compliance with study policies and procedures and regulatory requirements. At sites with inexperienced staff, the oversight of the principal investigator is even more important. Principal investigators should have regular meetings with study personnel to review the study protocols and participant enrollments to ensure that all personnel involved in the study are meeting the study requirements.

6.5 Background of Auditing Programs

The Food and Drug Administration (FDA) established guidelines for auditing and monitoring of clinical trials in 1977. The authority of the FDA to provide oversight of clinical investigations is defined by the Federal Food, Drug, and Cosmetic Act (FD&C Act) Section 505(i). The Code of Federal Regulations (CFR) provides regulations under Section 505(i) that describe sponsor responsibilities and clinical investigator responsibilities [4]. The guidelines for monitoring were updated in 1998 allowing more flexibility in the requirements for monitoring, establishing that the goal is high quality data. As long as the research entity can establish the quality of the data collected for the study, the monitoring is considered adequate. The previous requirements for onsite monitoring at regular intervals is being relaxed through increased use of technology to ensure data integrity. It is the expectation of the FDA that the investigator will comply with the Code of Federal Regulations with knowledge and understanding of clinical investigator regulations and responsibilities [5].

Prior to participating in any research program, the investigator must complete a Statement of Investigator Form 1572. Form 1572 requires information on the research resources at the investigators institutions, the planned protocol, and research affiliations of the investigator. The responsibilities of the investigator are outlined in the form. These are to personally conduct or supervise the study, to adhere to the protocol, ensure all staff participating in the study are trained and adhere to the study, inform subjects that drugs are being used for investigational purposes, ensure informed consent, ensure IRB review and approval of the initial protocol and all amendments, report adverse events as required by the sponsor and the FDA, and maintain adequate research records [5].

6.6 Goals of Auditing Programs

For many investigators, the audit process seems more like a test of their ability to complete research than a quality assurance and educational program for best research practices. However, the main objective of an audit program of any research organization is to verify the study data that would impact the primary outcomes of a study. In addition, the program is designed to keep the investigators informed regarding progress of the research program, educate programs regarding best practices, and clarify specific points regarding conduct of the trial.

6.7 Auditing of Clinical Trials

Auditing in large research organizations is generally governed by an audit committee. These are generally made up of research associates, physicians, and staff auditors. The primary function of the audit committee is to conduct the audits required by the group. The audit committees also produce an audit manual, provide education and training sessions for the research teams, provide feedback to the entire group regarding audit performance, identify components of the research study that are problematic from an audit perspective, and identify best and worst practices among the research sites.

Each audit will have three components of review: IRB and informed consent content, accountability of investigational agents and pharmacy operations, and patient case review. Each component has specific areas that must be reviewed and has specific requirements as outlined in federal guidelines. Areas that are not in compliance will receive either a major or minor deficiency. The total number of deficiencies in each component will be totaled to determine the overall evaluation of that component. Each component is scored as acceptable, acceptable needs follow-up, or unacceptable [6].

6.7.1 IRB and Informed Consent Content Review

This component evaluates the processes of the IRB and compliance with federal guidelines in the context of the approval of this study. The review of the IRB will include an evaluation of the regulatory binder. Maintenance of the regulatory binder is crucial to document compliance with this component. All correspondence with the IRB regarding the study should be kept in the regulatory binder. The IRB must review a sample informed consent, protocol/amendments and advertisements, written information provided to subjects, information about subject compensation, the Investigator's Brochure, and the investigator's current CV. The dates of IRB review and approval of amendments are noted. All amendments to the study must be

submitted, reviewed by the IRB, and approved within 90 days of the date of issue of the amendment. Throughout all phases of the study, the regulatory binder must be kept up to date to be considered in compliance with this component [7].

The consent content is also evaluated for compliance with federal guidelines. There are specific requirements for consent content. These can be identified by the major headings of a model consent form. When a model consent form is available, all components included in the model consent must be included in the institutional consent. The most common errors in writing the consent are leaving out some of the risks of the study and leaving out the appropriate contact information for the study and the study sponsors. The principal investigator must write an informed consent that contains the elements required by the FDA and includes the language required by both the sponsor and the institutional IRB [7].

The elements required for informed consent as required by the FDA are as follows: (1) A statement that this is research including duration of the participation and the study procedures that will be required. (2) A description of the possible risks and discomforts that the study subject may experience. (3) A description of the possible benefits that the subject may receive. (4) Alternative treatments and their potential benefits must be disclosed. (5) The method of protection of confidentiality of the study data must be described. A list of the organizations or individuals that will have access to the study data must be included. (6) A statement of possible compensation for injuries must be included if the study has more than minimal risk. (7) Contact information for the study sponsor and the principal investigator must be included. (8) The study must be described as voluntary and it must be stated that the subject will suffer no loss of benefits if they choose not to participate or if they withdraw from the study [8].

6.7.2 Accountability of Investigational Agents and Pharmacy Operations

All audits that include an investigational agent or agents supplied by the sponsor for the study are required to maintain accountability of those agents through the pharmacy. The audit will include an evaluation of the processes in the pharmacy and the methods of maintaining drug accountability [5]. Investigator responsibilities of investigational agents include maintaining control of the investigational agent with appropriate record keeping and retention of the agent. Investigational agents are kept in a research pharmacy. Each patient enrolled in the study will have a Drug Accountability Record Form (DARF) completed for tracking of the investigational drug. The drug must be tracked if it is not dispensed from the pharmacy but delivered to the patient by research staff. It is also important to maintain only the amount of drug necessary for the patient's enrolled in the study. If there is excess drug at a site, it should be returned to the sponsor as soon as it is determined that it is not needed. There are also requirements that investigational agents only be released to investigators that have a current Form 1572 and are an investigator in the trial.

In some centers, clinical personnel that are not a part of the research study may write prescriptions for the patient. It is important that the study drug must be prescribed by an investigator on the study. The pharmacy has specific requirements for maintaining the security of the investigational agent. This includes tracking of the agent, ensuring appropriate dispensing in compliance with the study, and the physical security of the pharmacy [5].

6.7.3 Patient Case Review

The final component is the patient case review. This component reviews all aspects of the study that involve a patient. This includes the informed consent and consent process, eligibility for participation in the study, treatment compliance with study procedures, disease outcome and response to treatment, adverse event reporting, and general data and management quality. This component requires the most preparation for the audit day. The cases to be audited will be provided by the audit team prior to the audit. Each research chart should be prepared for the auditors to facilitate the review of the information required. If the charts are not prepared, it is more difficult for the auditors to find the information needed and puts that information at risk of being noted as not available on the date of the audit. For each chart, the research staff should color code and label the elements of the chart that will be audited. For example, the source documents proving eligibility for the study should be flagged and labeled as eligibility criteria. This should be done for each major component of the patient case review. For all institutions with an electronic medical record, either the entire patient record will have to be printed or the auditors will have to have access to the electronic record. It is not sufficient to provide only the source documents for the study [7].

To complete the patient case review portion, a minimum of 10 % of patients enrolled will be reviewed. The sponsor will select those that have longer periods of follow-up and will try to audit charts of all investigators participating in the study. Each research chart and patient chart will be reviewed in its entirety. Each portion of this component will be reviewed and deficiencies noted [7].

The consent form and the consent process will be reviewed. A note documenting the consent process should be written by the person obtaining consent. It is important to note that the study subject had the opportunity to review the consent and that there was no coercion. The consent form will be reviewed to ensure that the correct version of the consent was used and that all required signatures were obtained. The IRB will date stamp the consent and the correct version of the consent should be used. Informed consent is required before any study procedures are completed. All signature lines on the consent form must be signed and dated including witness and principal investigator signatures if required by the IRB. Each individual providing a signature must date their own signature. It is unacceptable for study staff to date the consent for the subject. Some consent forms have lines for initials or checkboxes either at the bottom of each page or to participate in optional portions of the study.

These must be completed for the consent to be considered valid. It is unacceptable for the subject to mark through any portion of the consent form. The subject must agree to participate as the consent is written [8].

Eligibility is determined by reviewing source documentation of the inclusion and exclusion criteria. Each criterion must be confirmed with a source document. It is not adequate to use a signed case report form as a source document. If the inclusion criteria require that the patient remain on birth control for the duration of the study, there must be source documentation of this agreement. This can be in the form of a note in the medical record or a research document that is signed by the investigator. If source documents are not available at the time of the audit to confirm eligibility, then the patient will be considered ineligible until the documents are available. Preparation of the research charts by the research staff prior to the audit will ensure the availability of all the necessary source documents [6].

Treatment compliance refers to the sites adherence to the treatment or study protocols. This addresses whether the randomization was followed correctly and then the prescribed study treatment was implemented in accordance with the study protocol. This includes the timing of the treatments, proper dosing for pharmaceutical trials, and proper surgical techniques for surgical trials. For example, in a sentinel lymph node biopsy study was the surgical procedure completed following all protocol guidelines to ensure uniformity in the treatment arms. If a study is comparing sentinel lymph node to axillary dissection, the auditors will confirm that the actual surgical procedure completed was that required by the study based on review of the operative report and the descriptions of the procedures included in the protocol.

Each study will have specific endpoints that will provide a comparison of the research arms. This will include a measure of the disease outcome or response. Possible outcomes could be tumor response as measured by imaging or pathology studies, survival, survival without disease exacerbation, or morbidity resulting from the treatments. In oncology studies the most common outcome measure is survival or disease response. In vascular surgery studies, the outcome may be stroke for carotid interventions or recurrent ischemia for vascular bypass studies. For obesity studies, the outcome measure may be weight loss as a percent of body weight or total weight loss. The protocol will define the study procedures that are required to determine the primary and secondary outcomes. This may include imaging studies that must be done at specific time points after the intervention. The timing and completeness of the required follow-up studies will be documented as a part of the audit. Where an interpretation of the findings is required, a central review of those results may be required.

Adverse event reporting is required to determine the safety of the treatment arms. An adverse event is any change in the subject's health status while participating in the study. Some adverse events are related to the study intervention and others are not. For example, nausea and vomiting that occurs after taking a study medication is possibly related to the study drug and is reported as an adverse event and is attributed to the study intervention. If the patient falls and sustains a fracture, the fracture is an adverse event but is not related to the study intervention. There are standardized reporting methods for adverse events with standardized scoring of

those events. For studies of investigational devices, adverse events are reported using the Medical Device Reporting guidelines [9]. The study protocol will provide specific requirements for adverse event reporting that is in compliance with the federal regulations. All serious adverse events that are life threatening or fatal are required to be reported to the sponsor within 24 h and to the IRB. The study protocol will define expected adverse events; all others are unexpected adverse events. The expected adverse events of low severity do not have to be reported through the IRB. The unexpected events that are of low severity will be reported to the sponsor and the IRB but not in an expedited manner. All adverse events must be graded and then an attribution assigned. The attribution must be completed by a licensed clinician and will describe the relationship of the adverse event to the study interventions. Adverse events can be unrelated, probably caused by, or caused by the investigational agent. Those that are probably caused by or caused by the intervention will be reported to the sponsor and the IRB [7].

Adverse event reporting is a critical component of research studies to ensure the safety of the study participants. For studies of investigational agents or devices, a data safety and monitoring board may be required to follow the progress of the study to determine if the endpoints have been met or will not be met based on the interim analyses and to review the adverse events associated with the study. If automatic stopping rules for adverse events are reached, then the study will be closed by the data safety monitoring board for the study [10]. The final area that will be evaluated by the audit team is the general data management and quality. This area is a general assessment of the research process at the institution. The site will be evaluated for timeliness of data submission. Are study procedures completed and data submitted in a timely manner? Is the data reported generally accurate or are there errors in reporting? All studies required in the study calendar will have a window in which they should be completed as defined in the protocol. For example, if a diagnostic study is required on a yearly basis, there is usually a 1-month window for completion. The study results should then be provided to the study sponsor. There is usually a 3-month allowance to get the data to the sponsor. Data submitted beyond the 3 months is considered delinquent. A large number of delinquent submissions will result in a deficiency.

6.8 Exit Interview

At the conclusion of the audit, there will be an exit interview in which the audit team will inform the investigators of the findings of the audit. The exit interview must include the principal investigator or his designee, the audit team, members of the research team, and commonly includes a representative of the IRB. Specific deficiencies in each component will be defined and missing data will be reported. The exit interview is used by the audit team to not only report the findings of the audit but to answer questions from the audit team and to inform the investigators of progress in the study and potential changes that may be planned for the study.

If there are missing documents identified through the audit, the research team will have an opportunity to provide those documents to the audit team. If the research

team identified deficiencies while preparing the charts, the team can initiate changes in their research practice to correct those deficiencies. The deficiencies will still be recorded but could potentially be lesser instead of a major deficiency. The audit team will be able to assess the acceptability of the corrective action at the time of the audit.

After the audit, the audit team will generate an audit report. The report will list the deficiencies and generate a score for each component. Each deficiency will be scored as a major or lesser deficiency. The scoring follows standards outlined by the sponsor. The scores will be acceptable, acceptable needs follow-up, or unacceptable. All acceptable needs follow-up and unacceptable components will require a response to the audit in the form of a corrective action plan. The corrective action plan should focus on how the study team will alter their processes to prevent the deficiencies from being repeated on subsequent subjects. The corrective action plan will be reviewed by the sponsor and if acceptable, the site will proceed with the study. If the corrective action plan is unacceptable to the sponsor, the research site will have to revise the plan. For some sponsors, an unacceptable score will halt enrollment until the corrective action plan is acceptable.

Unacceptable audits will result in a reaudit in the next year or when adequate accrual to assess the site has occurred. For National Cancer Institute sponsored trials, two consecutive unacceptable audits in the same component will result in suspension of the site. The site would then have to demonstrate significant changes in the research structure and apply for reinstatement before participating in research trails again. In responding to unacceptable audits or acceptable needs follow-up, the site needs to look carefully at their processes and initiate changes as soon as the problems are identified [7].

6.9 Types of Site Visits Other than Auditing and Monitoring

Preparation for quality assurance reviews should start in the planning phase of the study. A well-trained, experienced research staff is crucial to the success of any study. A careful review of a proposed study should be completed by the research staff and the principal investigator to determine feasibility of completing the study at your institution while maintaining compliance with the study requirements at the institution. After deciding to participate in the study, several site visits between the sponsor and the study staff will occur. These are done to ensure that preenrollment requirements are met such as IRB approval and skills verification and to educate the staff as to protocol-specific procedures and requirements.

6.9.1 Pre-study Qualification Visit

The first visit is the pre-study qualification visit. At this visit, the sponsor will meet with the principal investigator, meet the study staff, and tour the facilities.

The purpose of this visit is to allow the sponsor to determine the ability of the site to complete the study. (NCI study visits) Preparation required for this visit other than ensuring availability of the staff and a meeting room is review of the protocol and assessment of interest and feasibility by the principal investigator. If the sponsor agrees to participation by the site, the next visit is the initiation visit [11].

6.9.2 Initiation Visit

The initiation visit occurs after the IRB approval is received and before the first patient is enrolled. The principal investigator, coinvestigators, and all of the study staff should attend this visit. The sponsor will review the regulatory documents prior to the visit. The visit is an extensive review of the protocol and study procedures. This review includes (1) the roles and responsibilities of study team personnel, (2) study protocols, (3) pharmacy operations and drug handling if a pharmaceutical agent is involved, (4) proper use of investigational devices if appropriate, (5) instruction in proper Case Report Form (CRF) completion, (6) record maintenance and management, (7) enrollment, consent, eligibility, and recruitment of subjects, and (8) adverse event reporting. Each of these topics will be reviewed in detail, and the study staff will have the opportunity to ask questions and clarify specific points in the protocol. This process is far more productive if the staff is already familiar with the protocol. When this visit is completed satisfactorily, the site will begin accruing patients [2].

6.9.3 FDA Audits/Inspections

FDA audits or inspections can be routine or for cause. Routine FDA audits are conducted to confirm the quality of the research program and the data submitted to the FDA. A routine FDA audit will be done on studies submitted to the FDA for a new drug application or a change in drug labeling. The sites selected for audit from a multicenter trial are those that have rapid accrual or are responsible for a high percentage of the patients enrolled. These are routine audits conducted to confirm the quality and reliability of the data submitted.

FDA audits for cause are conducted because there is some concern about the research practices at a site. An audit for cause may be requested because the principal investigator participates in a broad range of studies, completes studies outside of their specialty, enrollment is more rapid than expected from the site, or there is a larger than expected enrollment for a particular diagnosis. The FDA may also be notified of irregularities in the data either by the sponsor or by a review of the data submitted. Irregularities or inconsistencies in the data may result in an audit for cause.

Preparation for FDA audits is crucial. If an FDA audit is requested, the investigator should seek advice and assistance from the institutional research office and other experienced investigators. The site cannot refuse an FDA audit and should notify the IRB and the sponsor as soon as the request is received. These audits generally last 3–5 days and may include an audit of the institutional IRB.

The audit report from an FDA audit is FDA Form 483. The sponsor and the investigator will respond to the deficiencies identified in the report. The report and the response are the Establishment Inspection Report which is submitted to FDA headquarters. The final inspection findings of an FDA audit can be (1) no action indicated, (2) voluntary action indicated, and (3) official action indicated. No action indicated is similar to an acceptable routine audit and indicates the site is in compliance and no response is required. Voluntary action indicated requires a response from the site to address deficiencies identified at the site that have little impact on the study results. An official action is indicated when the research practices are objectionable and sanctions are required. This audit will require a response and a reaudit is likely [2].

Audits for cause can be conducted either through the FDA or through the Office of Human Research Protection's Division of Compliance Oversight. An audit for cause from either agency should be responded to in a similar manner and either can result in punitive actions against the investigator, the site, and the sponsor.

6.10 How to Learn More About Auditing

The most common errors that result in deficiencies identified in audits or monitoring visits are poor supervision and training of study staff, insufficient oversight by the principal investigator, inappropriate delegation of study tasks to unqualified persons, failure to adequately protect study subjects, and an overworked investigator or study staff. Principal investigators that know the study protocol well and supervise their staff with regular research meetings and review of the study data and procedures result in the most successful audits.

Most research organizations have educational opportunities available to research staff and principal investigators. Large multicenter trials will have study specific training programs for the staff which should be attended. Investigator meetings will be held to keep the principal investigators well informed of the progress of the study and of specific protocol issues that may have compliance problems. Each academic institution has required training for principal investigators and research staff. These are offered through the IRB and the research office.

In addition to educational opportunities, it is beneficial to attend an audit prior to your own audit. This has the most impact if you attend an audit as you establish your research program so that your processes are set early in the study to be in compliance with the audit program. In addition, if the organization offers auditor training, this is an excellent method to learn the audit process.

References

1. Weiss RB, Tuttle SS. Preparing for clinical trial data audits. J Oncol Pract. 2006;2(4):157–9.
2. Monitoring & Auditing of Clinical Trials. Developed by Center for Cancer Research, National Cancer Institute, NIH Supported by the CTN SIG Leadership Group.
3. Institutional Review Boards Frequently Asked Questions – Information Sheet Information Sheet - Guidance for Institutional Review Boards and Clinical Investigators website http://www.fda.gov/RegulatoryInformation/Guidances. Accessed 1 Mar 2013.
4. Investigator Responsibilities Regulation and Clinical Trials at http://www.fda.gov/downloads/Training/ClinicalInvestigatorTrainingCourse/UCM337271.pdf. Accessed on 21 Mar 2013.
5. Kleppinger CF. Investigator Responsibilities – Regulation and Clinical Trials. FDA'S 2012 Clinical Investigator Training Course. Division of Good Clinical Practice Compliance, Office of Scientific Investigations, Office of Compliance, CDER. Presented November 14, 2012 from www.fda.gov. Accessed on 21 Mar 2013.
6. Barrett B. Alliance Auditor Training. Presented at the Alliance Research Meeting, Chicago; 2013.
7. Clinical Trials Monitoring Branch, Cancer Therapy Evaluation Branch, and Division of Cancer Treatment and Diagnosis. Guidelines for Monitoring of Clinical Trials for Cooperative Groups, CCOP Research Bases, and the Clinical Trials Support Unit (CTSU). National Cancer Institute October 2006.
8. A Guide to Informed Consent – Information Sheet Guidance for Institutional Review Boards and Clinical Investigators Contents http://www.fda.gov/RegulatoryInformation/Guidances. Accessed on 19 Mar 2013.
9. Reporting Adverse Events (Medical Devices) at http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/PostmarketRequirements/ReportingAdverseEvents/default.htm. Accessed on 23 Mar 2013.
10. Guidance for Clinical Investigators, Sponsors, and IRBs Adverse Event Reporting to IRBs — Improving Human Subject Protection @ http://www.fda.gov/downloads/RegulatoryInformation/Guidances/UCM126572.pdf. Accessed on 23 Mar 2013.
11. Guidance for Industry E6 Good Clinical Practice: Consolidated Guidance. Additional copies are available from: the Drug Information Branch (HFD-210), Center for Drug Evaluation and Research (CDER), 5600 Fishers Lane, Rockville, MD 20857 (Tel) 301-827-4573 from http://www.fda.gov/cder/guidance/index.htm. Accessed on 21 Mar 2013.

Chapter 7
The Budget

Shuab Omer, Lorraine D. Cornwell, and Faisal G. Bakaeen

7.1 The Budget

The principal sponsors of biomedical research in the United States of America are as follows: (1) the federal government, (2) state and local governments, (3) private not-for-profit entities including foundations, and (4) industry [1, 2].

Research funding increased from $75.5 billion in 2003 to $101.1 billion in 2007; however, adjusted for inflation, it was only $90.2 billion. Similarly, adjusted for inflation, funding from 2003 to 2007 increased at a compound annual growth rate of only 3.4 % in comparison to an annual growth rate of 7.8 % from 1997 to 2003 [2]. Interestingly, in 2007, industry (58 %) was the largest funder [3], followed by the federal government (33 %). More recently, for fiscal year 2013, the National Institute of Health (NIH) has requested $ 30.860 billion for biomedical research, which is essentially unchanged from the enacted 2012 budget at $30.623 billion [18]. In 2011 there was an 18 % success rate for funding of R01 grants, which is in stark contrast with rates of 22 % in 2010, 25–32 % in 1993–2003, and 45–58 % in 1962–1966 [15–17]. The decrease in funding is considered to be due to a number of factors. Importantly, there has been an increase in the number of applications and an increase in current commitments to previously funded research projects, as evidenced by the fact that 75 % of the $15.8 billion that the NIH spent on extramural grants went to existing projects in 2010 [16].This dismal situation for NIH funding contrasts strikingly with all the new emerging avenues for research that are available now due to rapid advances in proteomics, genetic sequencing, stem cells, and other technological advances [15–17].

S. Omer, MD • L.D. Cornwell, MD • F.G. Bakaeen, MD (✉)
Cardiothoracic Surgery, Baylor College of Medicine, Michael E DeBakey Veterans Affairs Medical Center, Operative care line, Mail code: 112/ocl, 2002 Holcombe Boulevard, Houston, TX 77030, USA
e-mail: shuabomer@gmail.com; cornwel@bcm.edu; fbakaeen@bcm.edu

T.M. Pawlik, J.A. Sosa (eds.), *Success in Academic Surgery: Clinical Trials*, Success in Academic Surgery, DOI 10.1007/978-1-4471-4679-7_7,

From the preceding discussion, it is evident that there is no net increase in sponsorship for biomedical research and resources have not increased in comparison to the number of investigators applying for these grants. Thus, in today's economic climate, it is vital to not only have a scientifically sound project but also one that is economically viable. Having a realistic and thorough approach in formulating your budget is paramount. In this chapter, we will mainly discuss the process of budgeting for the two major sponsors of biomedical research, i.e., industry and NIH.

7.2 The Budget for NIH Sponsorship

The purpose of the budget is to present and justify all expenses required to achieve project aims and objectives. Formulating a budget can be challenging; however, the administrative officials at any institution and experienced peers can make this process much easier, especially for first-time investigators. It is important to figure out in advance the infrastructure of your institution regarding direct and indirect costs, fringe benefit rates, graduate stipend rates, facilities and administrative costs, etc., as these differ from institution to institution. There are certain logical steps you have to go through in order to submit the budget (Fig. 7.1). For multi-institutional study applications, a separate budget must be submitted for each participating site.

7.3 Complying with Federal Cost Principles

For a grant to be accepted by the NIH, not only should it be scientifically sound but should also comply with the governing cost principles. These cost principles are set forth in the NIH Grants Policy of allowable and unallowable costs [10]. For the NIH to approve your budget, the proposed costs charged to awards must be allowable, allocable, reasonable, necessary, and consistently applied regardless of the source of funds. There is a high likelihood of a proposal being rejected if these cost principles are not met [4].

7.3.1 The FOA (Funding Opportunity Announcement)

The FOA in addition to all the other information details the monetary limits on the types of expenses, like overall funding limits, construction allowed, and caps on travel expenses [4]. Before embarking on any project, carefully read the funding opportunity announcement for budget criteria and formulate your budget accordingly.

Do not under- or overestimate your budget, as it can adversely influence the chances of your proposal being accepted by suggesting to the reviewers that you do not understand the scope of the work involved. Reviewers keep in mind "the reasonable amount doctrine" to figure out whether the funds requested are justified by your aims and objectives.

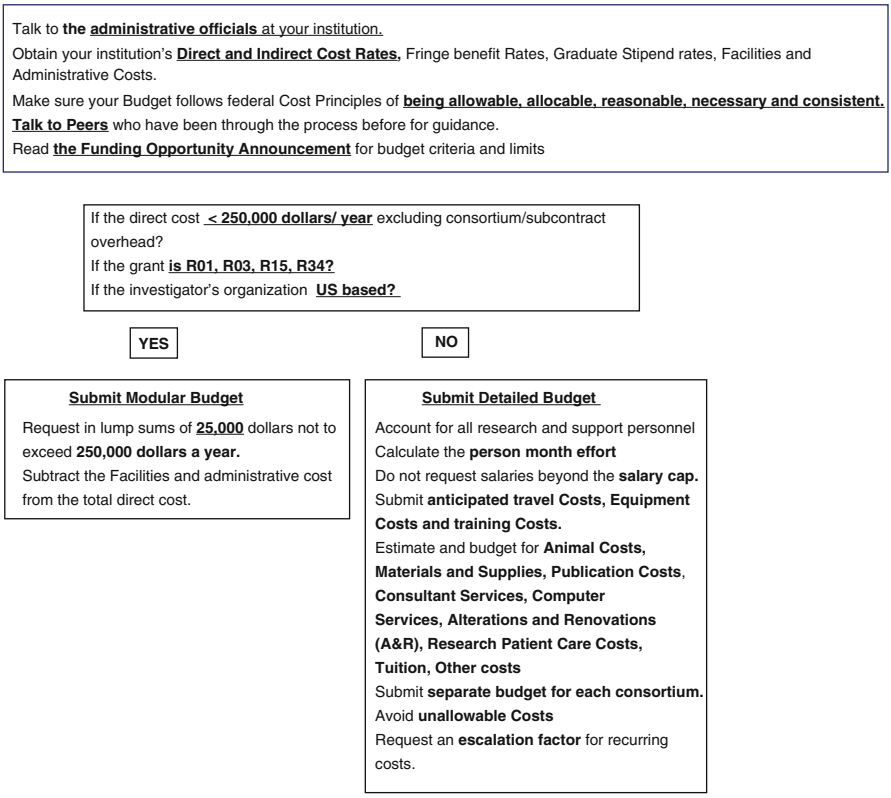


Fig. 7.1 Steps for submitting the NIH budget

7.4 Cost Sharing

Cost sharing implies charging a part of the cost of a sponsored project to a source other than the primary sponsor. In a university setup, this cost sharing contribution could be the cost and time of faculty members that commit to the project without charging the sponsor.

Sometimes a project requires cost sharing, as in large equipments awards. This is referred to as "required cost sharing." When cost sharing is desirable but not required, it is referred to as "voluntary cost sharing." This should be minimized whenever possible from your budget request. This cost sharing arrangement between your organization and the NIH does not normally impact the evaluation of your proposal [4].

7.4.1 Allowable Facilities and Administrative Costs (F&A Costs or Indirect Costs) and the Allowable Direct Costs

Direct Costs: These are costs that can be directly attributed to your project with ease and accuracy.

F&A Costs or Indirect Costs: These are costs associated with providing and maintaining the infrastructure that supports the research enterprise (buildings, maintenance, libraries, restrooms etc.); these cannot be easily identified with a specific program [4].

"Facilities" is defined as depreciation and use allowances, interest on debt associated with certain buildings, equipment and capital improvements, and operation and maintenance expenses [14]. "Administration" is defined as general administration and expenses, departmental and college administration, sponsored project administration, and all other expenditures not listed less than one of the subcategories of facilities [14].

F&A costs are determined in conjunction with auditors from the US Department of Health and Human Services for each institution. For profit organization, the F&A costs are negotiated by the Division of Cost Allocation (DCA), Division of Financial Advisory Services (DFAS) in the Office of Acquisition Management and Policy, and the NIH [4]. F&A costs are calculated by applying your organization's negotiated F&A rate to your direct cost base. In general for most institutions, the negotiated F&A rate will use a modified total direct cost (MTDC) base, which excludes items such as equipment, student tuition, research patient care costs, rent, and sub-recipient charges (after the first $25,000) [4].

It is also worth knowing that direct cost requests equal to or greater than $500,000 require prior approval from the NIH Institute/Center before application submission. For many SBIR/STTR(Small Business Innovation research/Small Business Technology Transfer) grantees, 40 % of modified total direct costs is a common F&A rate, although rates at organizations may vary.

7.4.2 Formats for NIH Budget Submission

The strategy for success is to propose simpler projects with lesser budgetary demands, as reviewers will scrutinize larger funding requests. Budget requests to the NIH can be submitted under two categories:

1. Modular budget
2. Detailed budget

For a new PI, a modular budget is preferable unless it cannot be avoided, as when the project requires >$250,000/year or you are based outside the United States of America. This becomes more evident on reviewing the FY 2010 data for competing R01 applications. The average application received roughly $290,000 in direct costs (not including the institution's overhead). About 76 % of new investigators used a modular budget. Applications requesting $250,000 or less in direct costs used a modular budget, and for non-new investigators, 66 % went the modular route [5].

7.5 Modular Budget

A modular budget format can be submitted if the direct cost is less than 250,000 dollars/year excluding consortium/subcontract overhead; the grant is R01, R03, R15, and R34; and the investigator's organization is US based.

The funds are requested in lump sums of $25,000. The numbers of modules requested are calculated by subtracting the overhead from the total direct cost and then rounding it to the nearest $25,000. Modular budgets do not automatically adjust for inflation for future years, so you have to plan the entire budget at the outset. Request the same number of modules annually, except for special needs such as equipment.

Even though not required when using a modular budget, it is worth creating a detailed budget for your own institution's use, including salaries, equipment, and supplies for funds requested. Even though these detailed expenses do not need to be submitted to the NIH, they are useful when calculating your overhead and for audits.

7.6 Detailed Budget

This budget format is used when the investigator's direct cost minus overhead is greater than 250, 000 dollars/year, the grant is not an R01, R03, R15, R21, or R34 grant. It is also used when the investigator's organization is not US based.

As the name implies, in this format, the investigators need to give detailed budgetary descriptions in the following areas: (1) research and support personnel involved; (2) equipment, travel, and training cost; (3) other direct costs; and (4) consortiums/subawards.

7.6.1 Research and Support Personnel

All research personnel from the investigator's organization involved in the project should be mentioned in the budget with their base salary and effort, irrespective of whether they are requesting salary support or not.

The funds requested for research and support personnel are requested in person months. Conversion of percentage of effort to person months is straightforward. This is done by multiplying the percentage of the personnel effort by the number of months of appointment.

For example, 10 % of a 10-month appointment = 1.0 person month ($10 \times .10 = 1.0$). Other issues to be addressed under this section are salary caps, fringe benefits, and senior/key personnel, which involves postdoctoral associates, graduate students, and other personnel.

Salary Cap: The NIH uses a salary cap to compensate the research and support personnel for your proposal. Requesting a salary above the salary cap will be counterproductive, as it results in a reduced total award amount. If in ensuing years the NIH increases the salary cap, the investigators can rebudget so that the personnel get paid as per the new cap [4].

The senior/key personnel who are devoting significant effort to the project should be mentioned. "Other significant contributors" who put meager effort should not be included. Examples of such common significant contributors include (1) CEOs of institutions providing overall leadership, but no direct scientific research contribution, and (2) mentors for K awardees, who provide advice and guidance to the candidate but do not directly work on the project. Consultants or associates who are not employed by the investigators' organization should not be appended as senior key personnel, but rather should be included in the section of the budget for consultants or in the category of the consortium/subaward budget page for collaborators.

Postdoctoral associates and graduate students should be entered as per the percentage of effort put in the budget justification section. When justifying people having the same job description such as "lab assistants and technicians," indicate the number of personnel involved with their role description, add their people months together, and add their requested salaries together. The salaries of secretaries and clerical staff are generally treated as overhead costs; if included as separate costs, their involvement should be directly and significantly related to the project [4].

7.6.2 Equipment, Travel, and Trainee Costs

Equipment is defined by the NIH as an item of property that has an acquisition cost of $5,000 or more (unless the organization has established lower levels), an expected service life of more than 1 year, be stand alone and function independently [4]. Sometimes replacement parts and fabricated equipment can be treated as exceptions to this standard definition. Generally, equipment is excluded from the facilities and administrative cost, so if you have something with a short service life (<1 year), even if it costs more than $5,000, you are better off appending it under the "supplies" category.

Routine equipment such as computers that will be used on other projects or for personal use should not be listed as a direct cost but should come out of the F&A costs, unless these items will be used solely for the actual conduct of the planned project.

Even when the application does not demand it, a price quote for new equipment, including price quotes in the budget proposal, can greatly help in the evaluation of the equipment cost to support the project.

Any time you request equipment that is costly, it is a wise strategy to first see if such equipment can be shared at your facility as this way you can cut down costs and have a better chance of success with the reviewers. In the event that the piece of equipment is vital and not available, then you will have to fully justify its need and also attest that it will be exclusively used for your project [5].

Your research project will require you or members of your team to travel. This has to be fully described in the budget request explaining the number of people traveling, dates, duration of your stay etc. It is again necessary that the travel has to be proximately related to the proposed research project. In the event that your institution lacks a specific policy for travel, then the US federal government policy in this matter can be adopted.

7.6.3 Budgeting for Other Direct Costs

These are (1) materials and supplies, (2) animal costs, (3) publication costs, (4) consultant services, (5) computer services, (6) alterations and renovations (A&R), (7) research patient care costs, (8) tuition, and (9) others.

7.6.3.1 Materials and Supplies

These include items that are expended or consumed in the conduct of the project, such as lab glassware, vials, chemicals, and reagents. Specify the amount for each item needed. However, categories that cost less than $1,000 do not have to be itemized.

7.6.3.2 Animal Costs

If your study involves live animals, then this can be included under "materials and supplies"; however, it is very convenient to include more specific details about how you calculated your estimate for animal costs. Include the number of animals you plan to use, the purchase price for the animals (if you need to purchase any), and your animal facility's per diem care rate, if available. Details become exceedingly helpful if your animal care costs are extraordinarily large or small. For example, if you plan to follow your animals for an abnormally long time period and do not include per diem rates, the reviewers may think you have budgeted too much for animal costs and may recommend a budget cut [4].

7.6.3.3 Publication Costs

The goal of research is to disseminate knowledge to bring about changes for the better. A research finding cannot have an impact unless it is published and reviewed. This could be a costly process, and thus, publication costs are important to be included in your proposal. In case of a new application, you can also delay publication costs until the later budget periods, once you have actually obtained data to share [4].

7.6.3.4 Consultant Services

Depending upon your project, you might require consultant support. For the NIH, consultants differ from consortiums in that they may provide advice, but should not be making decisions for the direction of the research [4]. They generally charge a fixed rate that includes both their direct and F&A costs; as a result, you do not need to report separate direct and F&A costs for consultants. However, you have to submit their travel cost estimates. Additionally, consultants are not subject to the salary cap restriction; however, any consultant fee should meet your institution's definition of "reasonableness" [4, 5].

7.6.3.5 Specialized Computer Services

This is separate from the general computer and professional support provided by your institution. This includes specialized supercomputer and software charges which, if needed, should be mentioned in your budget request.

7.6.3.6 Alterations and Renovations (A&R)

Setting up the infrastructure of your lab can be costly; simple things like making room for a new piece of equipment can strain the budget. Fortunately, you can request these charges in the budget under alterations and renovations. A&R does not include general maintenance projects, which are handled under overhead or projects exceeding $500,000, which are considered as "construction" projects. As expected, justify your expenses and itemize by category. If A&R costs are in excess of $300,000, further limitations apply, and additional documentation is required [4].

7.6.3.7 Research Patient Care Costs

This category includes costs for tests and procedures that are required only because the patient is participating in a research project and thus are not part of routine medical care. In general, only few NIH budgets request patient care expenses. In the event that your project involves both inpatient and outpatient expenses, you should mention the hospitals or clinics where care is to be rendered. You will also need to provide the details of how long you would be treating, number of patients enrolled, costs of treatment and diagnostic tests, etc. If both inpatient and outpatient costs are requested, the information for both of them are submitted separately [4].

7.6.3.8 Tuition

If you have graduate students working for your project, you will have to provide your school's tuition rates. Based on your institution's stipend and tuition rates, you

may at times have to budget less than your institution's full tuition rate in order to meet the graduate student compensation (equivalent to the National Research Service Award (NRSA) zero-level postdoctorate stipend level) [4].

7.6.3.9 Avoiding Unallowable Costs

The NIH has a list of questionable items under the NIH Grants Policy that are not allowed. It is advisable to identify and remove them up front because if the NIH identifies such an item, they will deduct it from your total award.

7.7 Consortiums/Subawards

Some research projects are undertaken as consortiums between the university/academic institution and businesses. In this respect, the NIH grants funding support via the Small Business Technology Transfer (STTR program) or Small Business Innovation Research (SBIR) [4, 7, 8]. When using the detailed budget format in this case, each consortium included must have a separate budget form filled out. In addition, regardless of what cost principles apply to the parent grantee, the consortium is held to the standards of their respective set of cost principles. Consortium F&A costs are not included as part of the direct cost base when determining whether the application can use the modular format (direct costs <$250,000 per year) or determining whether prior approval is needed to submit an application (direct costs $500,000 or more for any other year) [4].

If the consortium is a foreign institution or international organization, F&A for the consortium is limited to 8 %. If the consortium is with a for-profit entity, such as a small business, the organization must have a negotiated F&A rate before they can charge F&A costs. A default small business rate of 40 % is only applicable to SBIR (R43 &R44) and STTR (R41 & R42) applications. In addition, each consortium should provide a budget justification following their detailed budget. The justification should be in addition to the primary grantee's justification and address those items that specifically pertain to the consortium [4, 7, 8].

7.8 Predicting and Planning for the Future Years

The NIH does not expect your budget to foretell with accuracy what your expenses will be in a few years. However, they do expect an honest approximation of what your expenses might be. You can request an escalation factor for recurring costs in accordance with your institution's policy, depending on the NIH's budget appropriation. The NIH generally provides up to a 3 % escalation factor for recurring costs for each future year. Consistent with the FY 2009 appropriation, the FY

2008 average cost of competing grants is allowed to increase by 3 % over FY 2008 [9]. In general, NIH grantees are permitted to rebudget within and between budget categories to overcome unforeseen needs and to make other types of post-award changes. Some changes may be allowed at the grantee's discretion as long as they are within the limits established by the NIH. In other cases, the NIH needs prior written approval [10].

7.9 Budget for Industry-Sponsored Clinical Trials

Industry-sponsored clinical trials and research are pivotal contributors to biomedical research. There is roughly $6 billion in industry-generated money for clinical trials worldwide annually; out of this, $3.3 billion goes to the US investigators [11]. Because of the potential of great monetary benefits, approximately three quarters of funding for clinical drug trials in the USA is sponsored by industry rather than the NIH [12]. Industry-sponsored research thus represents the key supply of funding for an increasing number of clinical investigators [11, 13].

7.10 Understanding Ideological Differences Between You and the Sponsor

Sponsors and investigators view the proposed study very differently. The sponsors always will try to get the most out of the study by conducting it in the most expeditious and inexpensive fashion, which has the potential of undermining a lot of important details. The sponsors view the clinical trial contract as a fixed-price agreement. Thus, investigators are obliged to perform the task described in the contract, despite having exceeded the original proposed budget. Thus, successful budgeting for the performance of an industry-sponsored clinical trial requires a thorough understanding of all possible eventualities [6]. Keeping this in mind, we try to explain how best to plan such a budget.

7.11 Analysis of Direct and Indirect Costs

Similar to budgeting for NIH grants, it is again important to figure the direct and indirect costs at your institution, as it varies widely among institutions and countries (Tables 7.1 and 7.2). To thoroughly understand the direct and indirect costs associated with performing clinical research at a particular institution, the investigators should always first conduct an internal cost analysis independent of a sponsor's proposed budget.

Table 7.1 Direct costs usually incurred for industry-sponsored research

Staff salaries and benefits (investigators, nurses, consultants, etc.)
Training costs
IRB costs
Study initiation charges
Charges incurred with FDA audits and adverse outcome reporting
Data storage costs
Equipment and supplies
Mailing and shipping charges
Investigational device or drug preparation fees
Screen failure, delay, or dropout contingency charges
Scientific meeting and travel charges
Patient follow-up charges

Table 7.2 Indirect costs usually incurred for industry-sponsored research

Accounting charges
Building maintenance
Laboratory and office space maintenance and rent
Equipment wear and tear
Administrative costs
Utilities

In most research institutions, indirect costs are charged as a mandatory, fixed fee which is usually 20–40 % of the total direct cost [6]

7.11.1 Institutional Review Board (IRB) Charges

Before embarking on any investigational study in the USA, IRB approval is required. A lot of investigator time and effort is used up in this process and is easily overlooked or underestimated by the sponsor if not addressed in the budget. Furthermore, in the case of an industry-sponsored trial, it is not uncommon for a sponsor to amend the study protocol after initiation of a clinical trial, requiring additional IRB resubmissions. Also investigators are required to notify the IRB of the occurrence of any serious adverse events (SAEs) throughout the study (even if they occur outside the investigator's institution). This leads to additional expenses. Thus, it is important that the investigator requests all the time and labor costs related to IRB submissions, amendments, and reporting of adverse events [6].

7.11.2 Facilities and Administrative Cost (Institutional Overhead)

Similar to NIH-sponsored research, most major academic centers demand an institutional overhead of about 20–40 % of the total direct cost of the study. If an investigator

is at a site where indirect cost fees is not mandated by the institution, he/she should still budget for institutional overhead, as this money will be required to cover indirect costs such as rent, building maintenance, equipment depreciation, and basic utilities [6].

7.11.3 Laboratory Test Costs

All major institutions offer investigators a reduced research rate for the performance of in-house tests. However, these costs generally only cover test performance; thus, make sure to budget for all other additional charges incurred in this process, like collection of the sample, storage, and shipping [6].

7.11.4 Costs Associated with Preparation of the Investigational Device or Drug

Investigational drug and device studies may entail a variety of costs, including preparation, storage, dispensation, and accounting. Funds should thus be budgeted for training of the ancillary staff on preparation and handling of the new device [6].

7.11.5 Staff Salary and Training Charges

The major expense requested in a budget is for staff salary and training. The two points to remember are that this part of the budget is frequently over- or underestimated and that the charges vary from state to state. The staff involved range from consultants to investigators and nurses. The salaries requested should be commensurate to the amount of expertise and effort put in by each. The greater the complexity of a study, the greater the anticipated labor need, both in terms of the number of staff and salaried hours. It is also worthwhile to anticipate unforeseen charges, like collection of clinical data by staff at unusual hours when overtime rates may apply. Also anticipate and include charges such as those for device or drug preparation which could need costly consultant services and cause financial problems if not anticipated in advance [6].

It is also desirable to request the sponsor to pay the immediate costs of study initiation.

7.11.6 General Equipment and Supply Costs

This includes items such as phlebotomy supplies, centrifuges, freezers, computers, software, and copy/fax machines. In addition to direct equipment and supply costs, one should also budget for indirect costs such as equipment depreciation, extended service contracts, and secure patient record storage [6].

Table 7.3 Important issues to be addressed in the study contract for industry-sponsored trials

1. Specify the limits on the number of patients enrolled at your institution
2. Set up a payment schedule
3. Specify a start-up payment
4. Ensure charges for patient follow-up
5. Contingency in case of premature termination of study
6. Ensure funding for screen failures
7. Contingency funds for Food and Drug Administration (FDA) audits
8. Requesting charges for inflation adjustment for studies spanning a number of years

7.11.7 Patient Follow-Up

Since most clinical trials rely heavily on patient follow-up data, it is thus necessary to negotiate the cost of patient follow-up such as patient transportation, meals, and parking. Even if these expenses might be incurred later in the study, it is desirable to address them from the beginning [6].

7.12 Budgetary Considerations for the Study Contract

The study contract for an industry-sponsored clinical trial profoundly influences the chances of success of your project. Thus, the study contract should address several important issues (Table 7.3).

7.13 Initiation Charges

The study contract should also specify a certain sum of money to cover the investigator's immediate costs (e.g., staff training) while initiating the protocol, and it is not uncommon to ask the sponsor to pay the full price of one to three completed patients up front to cover the immediate costs of study initiation (i.e., money to be paid to the investigator before the first patient is enrolled) [6].

7.14 Backup Plan for Sudden Termination of Study, Delay, and Dropout

The study contract should also lay down specifics that should the study be terminated before enrollment of the first patient, the investigator's site will be compensated to cover start-up costs, payable immediately after an appropriate written notice. Similarly, it is useful to include compensation if the start-up of the study is

delayed due to unforeseen reasons or for screen failures. Screen failures are patients who are enrolled into a study but are subsequently barred from, dropout of, or are unable to participate in the study (e.g., a patient is enrolled in a study for Left Ventricular Assist Device (LVAD) implantation for destination therapy but ends up getting a heart transplant).

7.15 Ensuring Appropriate Number of Patients to Be Enrolled

Ensure that an appropriate number of subjects are being allowed to be enrolled at your institution as stated in the study contract. If the number expected is more than the anticipated volume at your institution, you will not be able to meet the contract expectations, and conversely if the number of subjects allowed are too small, then it is not worth your effort, as the amount of effort from an organizational standpoint is not very different up front whether you are enrolling 10 or 100 patients [6].

7.16 Reimbursement Timetable

It is very important to figure out in advance the reimbursement timetable as laid down in the study contract. This should address whether the payment is made at certain time intervals or at completion of study milestones. This is important because if a patient is lost to follow-up and as a result you cannot complete this milestone, then you need to address for this contingency. Furthermore, if your budgetary needs vary at different time intervals, then the compensation should reflect this.

It is also desirable to require that the sponsor be willing to cover the costs of returning a patient to the study site or sending a nurse to a patient's home for long-term follow-up if payment is dependent upon completion of specific milestones [6].

7.17 Audit Charges

The Food and Drug Administration does not audit the vast majority of studies. However investigators who have conducted decisive studies, acquire a large number of patients in a trial, or have participated in various phases of the same study are more likely to be selected for an audit [6]. Thus, it is advisable to have a clause in the contract that addresses such an event.

7.18 Inflation Adjustment for Studies That Span Over Years

It is also advisable to include an inflation adjustment to the study contract for studies anticipated to last longer than 1 year, as the cost of providing health care services is likely to increase over time.

Execution of a study contract without addressing the issues discussed in this chapter may significantly negatively impact the long-term budgetary goals of an otherwise well conceived study.

7.19 Summary

Planning your budget well and trying your best to foresee what your future needs would be is one of the most important parts of your research project. The most brilliant ideas might not come to any fruition if there is no monetary support. It is thus worthwhile to spend some extra effort in formulating all the details about your projected expenses. This will require talking to the administration at your institution and more importantly, to other peers who have been through the process before. Once you get the first few projects accomplished, then the process will become easier for you, as you will know how it works and even more importantly, the reviewers will take your proposals more seriously. Hopefully this discussion will get the reader better equipped with the challenges that he/she could face while preparing a budget proposal.

References

1. Campbell EG. The future of research funding in academic medicine. N Engl J Med. 2009;360(15):1482–3. nejm.org.
2. Dorsey ER, Roulet JD, Thompson JP. Funding of US biomedical research, 2003-2008. JAMA. 2010;303(2):137–43.
3. National Institutes of Health Office of Budget, Congressional Justification of the NIH fiscal year 2013 budget request. National Institutes of Health Web site. http://officeofbudget.od.nih.gov/. Accessed 12 Nov 2012.
4. US Department of Health and Human Services. Office of Extramural Research, NIH, Grants process Overview, Developing your Budget. National Institutes of Health Website. http://grants.nih.gov/grants/developing_budget.htm. Accessed on 12 Nov 2012.
5. Department of health and Human Resources, National Institutes of Allergy and Immunology, strategy For NIH funding. http://www.niaid.nih.gov/researchfunding/grant/strategy/Pages/3forms.aspx#j01. Accessed 17 Nov 2012.
6. Beal K, Dean J, Chen J. Budget negotiation for industry sponsored clinical trials. Anesth Analg. 2004;99:173–6.
7. National Institutes of Health, Research Funding, STTR Program. National Institutes of Health Website. http://www.nigms.nih.gov/research/mechanisms/STTR.htm. Accessed 26 Nov 2012.

8. National Institutes of Health, Research Funding, SBIR Program. National Institutes of Health Website. http://www.nigms.nih.gov/research/mechanisms/SBIR.htm. Accessed 26 Nov 2012.
9. NIH Fiscal Policy Grant Awards-FY-2009, NIH website. http://grants.nih.gov/grants/guide/notice-files/NOT-00-09-066.html. Accessed 27 Nov 2012.
10. NIH Office of Extramural Research, Grants and Funding. Changes in project and budget. NIH Website. http://grants.nih.gov/grants/policy/nihgps-2003/NIGPS-part7.htm.
11. Bodenheimer T. Uneasy alliance: clinical investigators and the pharmaceutical industry. N Engl J Med. 2000;342:1539–44.
12. Staropoli JF. MS. JAMA: funding and practice of biomedical research. JAMA. 2003;290:112.
13. Nathan DG. Clinical research: perceptions, reality, and proposed solutions: National Institutes of Health Director's Panel on Clinical Research. JAMA. 1998;280:1427–31.
14. The Ohio State University office of Research, Office of Sponsored programs. Proposal Budgets. Website: http://ops.osu.edu/development/. Accessed 28 Nov 2012.
15. Hormos R, Abkowitz JL, Keating A. Facing the NIH Funding crisis, How professional societies can help. Published Online: 26 Nov 2012. doi:10.1001/jama.2012.14846.
16. Office of Extramural Research, National Institutes of Health (NIH). What's behind the 2011 success rates? NIH website. 2012. http://nexus.od.nih.gov/all/201201/20/what%E2%80%99s-behind-the-2011-success-rates/. Accessed 31 Oct 2012.
17. Office of Budget, National Institutes of Health (NIH). Price Indexes. HIH website. http://officeofbudget.od.nih.gov/gbiPriceIndexes.html. Accessed 31 Oct 2012.
18. Ehrlich E. Engine Stalled: Sequestration's Impact on NIH and the Biomedical Research Enterprise. United for Medical Research website. http://wwwnitedformedicalresearch.com/wp-content/uploads/2012/07/UMR-Sequestration-Impact-on-NIH-2012.pdf. Accessed 14 Nov 2012.

Chapter 8
Regulatory Considerations in Human Subjects Research

H. Richard Alexander Jr., Edward Sausville, and Shannon Decker

8.1 Introduction

Biomedical research can be broadly defined as the systemic collection and analysis of data for the purposes of generating new knowledge that will relieve suffering and cure disease. Today, we understand that human subjects research must be conducted in compliance with federal statutes that are in place to ensure that all research activity is conducted ethically and follows the principles articulated in historical treatises such as the Nuremberg Code, the Helsinki Declaration, and the Belmont Report. For research involving drugs or devices, investigators also must comply with all US Food and Drug Administration (FDA) regulations relating to such research. This chapter will review the historical context under which these regulations were developed, provide an overview of the current regulatory requirements that must be met to perform human subjects research, and offer some practical considerations for new academic investigators.

H.R. Alexander Jr., MD (✉)
Division of General and Oncologic Surgery,
The Department of Surgery, Marlene and Stewart Greenebaum Cancer Center,
University of Maryland School of Maryland, Baltimore, MA, USA

Division of General and Oncologic Surgery, Department of Surgery, University of Maryland School of Medicine, 22 South Greene Street, Baltimore, MA 21201, USA
e-mail: hralexander@smail.umaryland.edu

E. Sausville, MD, PhD • S. Decker, MPH, JD
Division of Medical Oncology, The Department of Medicine, Marlene and Stewart Greenebaum Cancer Center, University of Maryland School of Maryland,
Baltimore, MA, USA

T.M. Pawlik, J.A. Sosa (eds.), *Success in Academic Surgery: Clinical Trials*,
Success in Academic Surgery, DOI 10.1007/978-1-4471-4679-7_8,

8.2 Historical Perspectives

Biomedical research on human subjects has been performed for centuries; perhaps one of the most celebrated examples of a prospectively conducted controlled clinical trial was the one performed by the Scottish physician James Lind aboard the HMS Salisbury in 1747. At that time scurvy, a condition that is now known to be caused by vitamin C deficiency, was common in sailors who were at sea for extended periods of time and who did not have access to fresh fruits and vegetables, a major source of ascorbic acid. Mr. Lind had a strong interest in improving the health of British sailors and conducted a study to evaluate various remedies for this common but poorly understood condition. He assigned 12 sailors with the typical signs and symptoms of advanced scurvy to six groups of two each. They all received the same diet but, in addition, each cohort received a daily regimen of either a quart of cider, a teaspoon of sulfuric acid, six spoonfuls of vinegar, a cup of seawater, a drink of barley water, or two oranges and a lemon. The experiment lasted for 1 week until they ran out of fruit, but by that time, the fortunate sailors in the last group had demonstrated dramatic improvement in symptoms. While this study exemplifies how clinical research can make meaningful and even dramatic discoveries that can eliminate suffering and cure disease, there are other unfortunate examples of clinical research that were conducted under unethical and even appalling circumstances that have provided the impetus for the development of our current regulatory infrastructure.

Beginning in 1932, the US Public Health Service conducted a 40-year clinical study to characterize the natural history of untreated syphilis in 600 indigent poorly educated rural black men in Macon County, Alabama. Study participants were enticed to participate by being told they would receive free health care from the US government, meals, and free burial insurance. Investigators never told the study participants that they had syphilis. Perhaps the most significant breach of ethical conduct relates to the fact that researchers knowingly failed to treat study participants for their syphilis even after the validation in the early 1940s that penicillin was an effective cure for this condition. Even after penicillin had become the widely accepted standard of care for syphilis in the late 1940s, study investigators continued to withhold treatment and actively prevented study participants from receiving it from other health facilities in the area.

In addition to that study, from 1946 to 1948, the US Public Health Service in collaboration with Guatemalan health authorities conducted a study in which prisoners, soldiers, and mentally ill patients in that country were deliberately infected with syphilis and other sexually transmitted diseases without obtaining informed consent. Although subjects were treated for their condition once infected, there was never any documentation of cure (www.wikipedia.org/wiki/Syphilis_experiments_in_Guatemala, accessed 12-28-2012).

8.2.1 The Nuremberg Code and Helsinki Declaration

Following World War II, the Nuremberg Military Tribunals were a series of 12 US military tribunals to prosecute war crimes against members of the leadership of

Table 8.1 Points of the Nuremberg Code (italics added)

1. The *voluntary consent of the human subject is absolutely essential*. The person involved should have legal capacity to give consent; should be so situated as to be able to exercise free power of choice, without the intervention of any element of force, fraud, deceit, duress, over-reaching, or other ulterior form of constraint or coercion; and should have sufficient knowledge and comprehension of the elements of the subject matter involved as to enable him/her *to make an understanding and enlightened decision*. This latter element requires that before the acceptance of an affirmative decision by the experimental subject there should be made known to him the nature, duration, and purpose of the experiment; the method and means by which it is to be conducted; *all inconveniences and hazards reasonable to be expected*; and the effects upon his health or person which may possibly come from his participation in the experiment. The *duty and responsibility for ascertaining the quality of the consent rests upon each individual who initiates, directs or engages in the experiment*. It is a personal duty and responsibility which may not be delegated to another with impunity.
2. The *experiment should be such as to yield fruitful results for the good of society*, unprocurable by other methods or means of study, and not random and unnecessary in nature.
3. The experiment should be so designed and based on the results of animal experimentation and knowledge of the natural history of the disease or other problem under study that *the anticipated results will justify the performance of the experiment*.
4. The experiment should be so conducted as to *avoid all unnecessary physical and mental suffering and injury*.
5. *No experiment should be conducted where there is a prior reason to believe that death or disabling injury will occur*; except, perhaps, in those experiments where the experimental physicians also serve as subjects.
6. The *degree of risk to be taken should never exceed that determined by the humanitarian importance of the problem* to be solved by the experiment.
7. *Proper preparations should be made and adequate facilities provided to protect the experimental subject* against even remote possibilities of injury, disability, or death.
8. The *experiment should be conducted only by scientifically qualified persons*. The highest degree of skill and care should be required through all stages of the experiment of those who conduct or engage in the experiment.
9. During the course of the experiment *the human subject should be at liberty to bring the experiment to an end* if he has reached the physical or mental state where continuation of the experiment seems to him to be impossible.
10. During the course of the experiment *the scientist in charge must be prepared to terminate the experiment at any stage*, if he has probable cause to believe, in the exercise of the good faith, superior skill and careful judgment required of him that a continuation of the experiment is likely to result in injury, disability, or death to the experimental subject.

Nazi Germany that were perpetrated against prisoners of war including brutal medical experimentation. The tribunals were held in the Palace of Justice, Nuremberg, after World War II from 1946 to 1949. The Nuremberg Code is a set of research ethics that were derived as part of the verdicts of the trials and define broad guiding principles for human experimentation. The Nuremberg Code includes such principles as informed consent and absence of coercion, properly formulated scientific experimentation, and beneficence towards experiment participants. The 10 points of the Nuremberg Code are highlighted in Table 8.1 (history.nih.gov/research/downloads/nuremberg.pdf, accessed 12-27-2012).

Table 8.2 Principles of the Belmont Report

1. Respect for persons: Protecting the autonomy of all people (research subjects), treating them with courtesy and respect, and providing informed consent. Researchers must be truthful and conduct no deception
2. Beneficence: The philosophy of "do no harm" while maximizing benefits for the research project and minimizing risks to the research subjects
3. Justice: Ensuring reasonable, nonexploitative, and well-considered procedures are administered fairly—the fair distribution of costs and benefits to *potential* research participants—and equally

Subsequently, in the 1960s the Declaration of Helsinki was developed by the World Medical Association as a set of ethical principles regarding human experimentation for the medical community. It is widely regarded as the cornerstone document of human research ethics. It is not a legally binding instrument, but has been used as the basis for legal statues and regulations overseeing human subjects research in numerous countries including the United States. The Declaration was originally adopted in June 1964 in Helsinki, Finland, and has since undergone multiple revisions.

Prior to the 1947 Nuremberg Code there were no broadly established principles that addressed the ethical aspects of human research. The Helsinki Declaration was based on the principles first stated in the Nuremberg Code, with some modifications. For example, the Declaration promoted a broader definition of the need for informed consent from "absolutely essential" under Nuremberg to "if at all possible"; research was allowed without consent where a proxy consent, such as a legal guardian, was available.

8.2.2 The Belmont Report

As a result of the Tuskegee Study and influenced by the tenants of the Helsinki Declaration, the National Research Act was signed into law by Congress on July 12, 1974. The Act authorized the creation of the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. One of the charges to the Commission was to identify the basic ethical principles that should underlie the conduct of biomedical and behavioral research involving human subjects and to develop guidelines that should be followed to assure that such research is conducted in accordance with those principles. The final report was issued in 1978 and was entitled *The Belmont Report*: *Ethical Principles and Guidelines for the Protection of Human Subjects of Research*, *Report of the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research*. It is commonly referred to as "the Belmont Report" and takes its name from the Belmont Conference Center in Elkridge, Maryland, where it was drafted (www.wikipedia.org/wiki/Belmont_Report, retrieved 12-29-2012). The Belmont Report defines three principles described in Table 8.2 (http://ohsr.od.nih.gov/guidelines/belmont.html).

8.3 Regulation of Human Subjects Research

8.3.1 The Code of Federal Regulations and Federalwide Assurance for the Protection of Human Subjects

In 1991, 14 federal departments and agencies joined HHS in adopting a uniform set of regulations for the protection of human subjects, identical to subpart A of 45 CFR part 46 of the HHS regulations. This uniform set of statutes constitutes the Federal Policy for the Protection of Human Subjects, informally known as the "Common Rule." The Office for Human Research Protections (OHRP) was also established under the Assistant Secretary for Health in the Department of HHS; it is responsible for the protection of the rights, welfare, and well-being of research subjects in research conducted or supported by the US Department of Health and Human Services (www.hhs.gov/ohrp/about/index.html). The OHRP principally interacts with biomedical research institutions to ensure compliance with HHS regulations as described in Title 45, Part 46, Code of Federal Regulations (45 CFR 46). The Division of Education and Development provides guidance to individuals and institutions conducting HHS-supported human subjects research. The Division of Policy and Assurances administers the Federalwide Assurance (FWA) of compliance and registration of institutional review boards.

The FWA is applicable to any institution that is engaged in human subjects research that is conducted or supported by any US federal department or agency that has adopted the Common Rule. The institution must renew its FWA every 5 years, even if no changes have occurred, in order to maintain an active FWA. There are rare circumstances under which human subjects research is exempt from the Common Rule but almost all research conducted at academic biomedical research institutions is covered under the FWA. The FWA number may be needed for grant applications that involve human subjects research, and the number should be available to an individual researcher from the institutional Human Research Protection Office (HRPO). A description of the elements of an FWA is listed at www.hhs.gov/ohrp/assurances/assurances/filasurt.html. The FWA includes a statement from the institution that all human subjects research will be conducted ethically and that the rights and welfare of human research subjects will be protected. The principles are generally adopted from the Declaration of Helsinki and the Belmont Report. The institution must provide a description of procedures to ensure prompt reporting of any deviations of principles and policies to the institutional review board, the US federal department or agency conducting or supporting the research and OHRP.

8.3.2 Institutional Review Boards

Under the FWA, an institution must constitute an Institutional Review Board (IRB) to review, approve, and provide oversight of human subjects research. The institution may be required to provide its written procedures regarding human subjects

Table 8.3 Conditions under which protected health information may be collected without informed consent

When the collection and use of PHI will pose no more than "minimal risk" to the individual and the investigators provide a plan to ensure that the information is properly collected, stored, analyzed, and ultimately destroyed
When the research could not practicably be conducted without the waiver
The researchers show that the use of PHI is essential for the success of the research

research to the OHRP or any US federal department or agency conducting or supporting research to which the FWA applies. Based on the 45 CFR 46, an IRB is required to (1) conduct initial and continuing annual reviews of research and report its findings to the investigator and the institution, (2) determine which projects require review more often than annually and which need verification from sources other than the investigator that no material changes have occurred since the previous IRB review, (3) ensure prompt reporting to the IRB of proposed changes to any research activity, and (4) ensure that proposed changes in approved research, during the period for which IRB approval has already been given, may not be initiated without IRB review and approval (except when necessary to eliminate apparent immediate hazards to the subjects). The last point is important in that it means that any investigator may deviate from the approved research plan if, in his or her judgment, that deviation is essential to reduce an immediate risk to the subject.

8.3.3 The Health Insurance Portability and Accountability Act (HIPAA)

In 1996, the Health Insurance Portability and Accountability Act (HIPAA), which codifies rules regarding the security of protected health information (PHI) by medical practitioners and "covered entities," added another layer of regulatory requirements to human subjects research. HIPAA compliance is required not only in everyday medical practice but in clinical research as well (www.hhs.gov/ocr/privacy/, accessed 12-31-2012). Safeguards to ensure confidentiality of PHI that is collected as part of a research activity are required by the institutional HRPO and are the investigator's responsibility. The Privacy Rule is designed to protect an individual's identifiable health information while allowing researchers to have access to vital medical information that is necessary to their research activities. Currently, most research involving human subjects operates under the Common Rule and/or the Food and Drug Administration's (FDA) human subject protection regulations. In clinical research activities, a part of the informed consent process must include a discussion of the investigator's intention to collect PHI as an integral part of the study. IRBs will require an explicit plan that describes how the PHI will be collected, stored, analyzed, and ultimately destroyed once the research activity is completed. Under certain circumstances, a waiver may be requested from the IRB to collect PHI without explicit informed consent (Table 8.3).

Once requirements for investigator training have been completed, a clinical trial may be submitted to the IRB. Most AHCs have a two-tiered system of review; the initial review is typically performed at a departmental or center level before going to the IRB. Once a study has been approved, an investigator is responsible for its timely completion in compliance with the institutional requirements. These include appropriate screening and review of eligibility, informed consent, accurate and complete follow through of protocol design, accurate and timely completion of case report or study forms, timely submission of annual reviews, and formal study closure at completion.

8.3.4 Food and Drug Administration and Clinical Research

Just as there were abuses in clinical research experiments that led to the development of the modern regulatory structure in which clinical research is conducted, both deliberate and reckless tragedies with drugs and devices occurred throughout the early parts of the twentieth century, from which evolved the Food and Drug Administration's (FDA's) modern role in the regulation of drugs and devices. Although it has been amended more than a hundred times since its passage, the 1938 Federal Food, Drug, and Cosmetic Act underpins the current regulations. Where clinical research involves drugs or devices, both FDA regulations and HHS regulations on human subjects research must be followed. While in many cases, the FDA and OHRP require the investigator to meet the same standards to satisfy both sets of regulations, an investigator needs to be aware that differences exist, and the investigator is still responsible for following both sets of regulations.

The US FDA requires an investigational new drug application (IND) or investigational device exemption (IDE) from an investigator or a sponsor (industry or collaborative group) under the following conditions:

- If the research is designed to establish a new marketing indication
- If the research is designed to establish a new dose or route of administration
- If the research is designed to define a new patient population not currently identified
- Significant change in the promotion of an approved drug

The purpose of an IND or IDE is to assure that research subjects will not be subjected to unreasonable risk. Within the FDA the Center for Drug Evaluation and Research (CDER) is the entity that is responsible for oversight of new drug evaluation prior to marketing. The Center for Devices and Radiological Health (CDRH) is responsible for regulating organizations or entities that manufacture, repackage, relabel, or import medical devices sold in the United States. If an investigator is unsure as to whether or not an IND or IDE is required, the institutional HRPO or FDA should be consulted. The FDA's regulations regarding the conduct of clinical research are defined in the CFR Title 21 and are in place to ensure compliance with Good Clinical Practices (GCPs) (Table 8.4). The FDA has oversight of clinical

Table 8.4 Elements of study conduct according to Good Clinical Practice

IRB and ethical study conduct
Informed consent
Eligibility criteria defined and changed only with IRB approval mechanism
Qualifications of investigators: Physician should be PI and assume responsibility for clinical consequences and follow-up of consequences of drug, biologic agent, or device action
Use of investigational materials under an IND or equivalent certification of Good Manufacturing Practice (GMP)
Adequate resources including space for conduct of trial and ancillary personnel to assume responsibility for research as opposed clinical care demands
Compliance with protocol: Deviations permissible ONLY to eliminate immediate prospect of harm to subject or relate to logistical or administrative aspects of trial
Compliance with protocol: Auditing system to assure proper entry of data into paper or electronic case report forms vs. original source documents. This is referred to as a quality control system
Define independent data safety monitoring committee or process for all greater than minimal risk trials
Definition of an adverse event collection and reporting system, with definition of critical events reportable to sponsor, and IRB outside of routine (i.e., at least annual) reporting periods
Accountability for receipt and use of investigational products (drugs/biologicals) according to specifications appropriate for investigational agent (contained in its IND application) or manual of use (devices)

studies that involve an IND or IDE and actively monitors to ensure compliance with study design. Informed consent violations continue to be the most serious violation identified by the FDA.

The concepts associated with GCPs include the ethical considerations introduced above, along with several points of detail in how to design and conduct protocols. The International Conference on Harmonization reflecting jointly recognized standards by US, European, Japanese, and other regulatory agencies are embodied in a Guidance for Industry (E6 Good Clinical Practice: Consolidated Guidance) available through http://www.fda.gov. Although it is true that exploratory protocols by academic researchers are frequently not at a stage where a clinical trial outcome will result in a regulatory approval of a drug, biological, or device, many institutions require all treatment protocols to follow GCP guidelines. Also, even in the absence of institutional requirement for GCPs, in the event the researcher is audited by the FDA or by a prospective corporate partner for further investment in the idea, clear evidence of following GCPs will increase enthusiasm for the credibility of the findings, and in the case of an FDA audit, avoid a publicly available citation for failure to follow GCPs.

8.3.5 Good Manufacturing Practices

Surgical physician-scientists frequently are focused on ultimately applying a local treatment or delivery approaches involving a drug or biological agent (e.g., virus, DNA construct, engineered cell). Frequently these materials may be derived from

an academic laboratory. In order to comply with GCPs, such materials that are not already approved and available for clinical use require an IND (Investigational New Drug) application. Materials used under an IND must be manufactured under Good Manufacturing Practices (GMP).

GMP includes the process of providing a complete description of the agent. In the case of a drug, this is usually a chemically defined structure. In the case of biologicals, focus on purity and potency of the product is key where molecular definition is not possible. The manufacturing process must assure that there is no reasonable likelihood of contamination of the final product with chemically or biologically injurious contaminants and that a process for serial monitoring of stability of the investigational agent under the proposed conditions of storage is in place.

Of particular relevance to biological agents is a characterization of cell lines that may have been used in their manufacture with respect to viral or other microbiological contamination; the relation of the cells expanded during manufacture to a master cell bank usually derived from a well-characterized (by sequence) nucleic acid construct-transfected or antigen producing cell type. In particular, if allogeneic cells are to be introduced into a human host, strategies to prevent their replication or production of recombinant infectious agents must be considered. This includes the development of "release criteria" for use in human clinical activities of a cellular product. In the event the cellular product is derived by culture from the patient's own autologous cells, careful definition of the conditions of expansion of the cellular product, monitoring for infectious agents appearing during processing, and time between completion of processing and use should be clearly delineated. All aspects related to production of the biological product should be described in a way that allows audits to assure quality of product use, and in the event of an adverse event, facilitate review of product integrity as it may be related to the clinical experience.

8.3.6 Investigator Responsibilities

The regulations described above place responsibilities on many parties: a drug or device manufacturer (to produce the drug or device under study in compliance with regulations), the sponsor of the study who may or may not be the manufacturer (to oversee the proper conduct of a study), an investigator's institution (to maintain an FWA and assure that either an internal or external IRB oversees the study), the IRB (to review the research both initially and while ongoing), and finally on the investigator. An investigator must first and foremost comply with institutional requirements that ensure there is adequate training regarding an investigator's responsibilities and the principles of human subjects research. In most academic health centers, those requirements are described within the institutional HRPO's website. Many centers use the Collaborative Institutional Training Initiative (CITI) program to fulfill investigator training. The Collaborative Institutional Training Initiative (CITI) was established in March 2000 as a collaboration between the

University of Miami and the Fred Hutchinson Cancer Research Center to develop a web-based training program in human research protection. Currently, content for the program comes from ten centers, and it includes numerous modules on various dimensions of human research. New investigators are required to pass the basic modules, and then depending on the nature of the investigator's research activities, additional modules may be required (www.citiprogram.org, accessed 12-31-2012).

Once an investigator has received the training appropriate to conduct research, the investigator then takes on additional responsibilities related to the specific research. The investigator must obtain all institutional approvals for human subjects research, then must obtain IRB approval for the research. If the investigator is writing the clinical trial protocol or manufacturing the drug or device, the investigator then also has sponsor and/or manufacturer requirements to comply with all applicable HHS and FDA regulations.

Adverse event data collection is probably one of the most time-consuming and confusing parts of any human subjects research project. The sponsor, the FDA, the NIH or other funding agency, the IRB, and the Data & Safety Monitoring Board for the study are all likely to have slightly different reporting requirements, and the investigator has to comply with all of them.

8.3.7 Practical Considerations for New Investigators

With the time pressure on new academic faculty members, navigating the many levels of approval necessary prior to initiating research can be daunting. An investigator should find out before even beginning to work on a research project what approvals are needed at the specific institution and what deadlines have to be met to obtain those approvals. An investigator should also determine what resources the institution has to help obtain those approvals, such as regulatory coordinators, IRB staff members, and/or investigator training specific to the institution.

Conducting clinical research requires an investigator to operate under a stiffer set of rules than necessary when managing only the clinical care of a patient. However, many new investigators confuse what is acceptable in clinical practice with what is required for research. For instance, unless special permission is obtained from an IRB in advance, informed consent for research may only be obtained in writing, in a language understandable to the participant, by the participant. This means that unlike clinical practice, without explicit permission from an IRB, oral consent is not adequate; a translator may not translate consent orally for a participant; and the next of kin or power of attorney may not give consent. Being cited by any regulatory body for failing to obtain valid informed consent is a serious violation and can result in the data being ruled unusable.

In clinical practice a physician has greater leeway to make substitutions and adjust to the realities of the moment. If a protocol, however, specifies a particular gauge needle, a particular supportive care medicine or a specific type of tube for a blood draw, any substitution is a protocol deviation even if the substitution has no

clinical or scientific relevance whatsoever, and an auditor will still have to cite the investigator for the deviation. Many problems can be avoided by thinking carefully about what details to insert into a protocol. An investigator should provide only those details that actually impact the conduct of the research. If certain antiemetics should be avoided due to interactions with a drug used in the study, by all means the investigator should prohibit them, but if any antiemetic will do, then specifying the particular one only creates the potential for deviations.

Many new investigators have difficulty assessing precisely when their review of patient records crosses the line into research that requires IRB approval. Even when patients are the investigator's patients or his/her practice group's patients, an investigator can still be cited for conducting human subjects research without IRB approval if the review of records extends beyond that needed for clinical care. A new investigator therefore needs to find out up front the institution's definition of a case report or case series and when the institution's IRB requires approval before the investigator can publish. Also, valid internal quality control projects can easily transition into research requiring IRB approval. For example, an investigator may be reviewing all cases in the practice group to determine if standardizing SOPs for the group could result in lower costs. In the midst of doing so, the investigator may realize that applying those SOPs results in better outcomes for the patients, and those results might be useful to physicians at other institutions. At this point, the investigator should submit an IRB application so that the data can be further analyzed and the results can ultimately be published.

8.4 Conclusions

Regulatory requirements in human subjects research have evolved from guiding principles designed to protect the rights and welfare of the research subject. However, they also serve to protect the investigator and the integrity of clinical research. In this way, clinical research has the best likelihood of providing meaningful new discoveries that will relieve suffering and cure illness.

Suggested Reading

1. Brown SR. Scurvy: how a surgeon, a mariner, and a gentleman solved the greatest medical mystery of the age of sail. New York: St. Martin's Press; 2003. ISBN 0-312-31391-8.
2. Rockwell DH, Yobs AR, et al. The Tuskegee study of untreated syphilis; the 30th year of observation. Arch Intern Med. 1964;114:792–8. PMID 14211593.
3. Jones J. Bad blood: the Tuskegee syphilis experiment. New York: Free Press; 1981. ISBN 0-02-916676-4.
4. www.wikipedia.org/wiki/Syphilis_experiments_in_Guatemala.
5. Goodyear MDE, Krleza-Jeric K, Lemmens T. "The Declaration of Helsinki". BMJ. 2007;335(7621): 624–5. doi:10.1136/bmj.39339.610000.BE. doi:10.1136%2Fbmj.39339.610000.BE.
6. www.fda.gov.
7. www.citiprogram.org.

Chapter 9
Publishing a Clinical Trial

Babak J. Orandi, Julie A. Freischlag, and Mahmoud Malas

9.1 Introduction

The publication of investigative efforts is a critical metric for promotion of tenure-track faculty, making it a necessity for success in an academic surgical career. More importantly, there are ethical imperatives to publish the results of clinical trials involving human subjects. Published results allow for the dissemination of findings that may improve patient care and outcomes. Patients who participate in clinical trials often do so with little realistic possibility of benefit from involvement with the trial, particularly in phase I and II trials [1, 2]. Their heroic sacrifice merits recognition in the form of publication of the data to which they contributed for the benefit of others. Given the current era of research budget constraints, it is all the more important to prevent duplicative investigative efforts by publishing work that has already been done. Failing to publish trial results contributes to bias in the overall body of literature on a given topic [3, 4]. This is particularly true for negative results [5–7]. This makes subsequent meta-analyses and systematic reviews difficult to perform and less informative [8, 9]. Finally, the knowledge gained from publishing the results of the trial do not necessarily relate to the results themselves, but rather to the methodological approaches and challenges that may be of interest to others in the same or similar lines of research [10].

B.J. Orandi, MD, MSc (✉) • J.A. Freischlag, MD
Department of Surgery, Johns Hopkins Hospital, Baltimore, MD, USA
e-mail: borandi1@jhmi.edu; jfreisc1@jhmi.edu

M. Malas, MD, MHS
Department of Vascular and Endovascular Surgery,
Johns Hopkins Bayview Medical Center, Baltimore, MD, USA
e-mail: mbmalas@jhmi.edu

T.M. Pawlik, J.A. Sosa (eds.), *Success in Academic Surgery: Clinical Trials*,
Success in Academic Surgery, DOI 10.1007/978-1-4471-4679-7_9,

9.2 Before the Trial Begins

Much of the work of publishing a clinical trial begins long before the writing phase. Requisite to a good write-up is good study design. Though beyond the scope of this chapter, the importance of sound study design cannot be overemphasized.

All human subjects trials involving interventions with drugs, biologics, and devices, with few and very limited exceptions, must be registered with ClinicalTrials.gov [11]. Section 801 of the Food and Drug Administration Amendments Act provides details about exceptions to the mandatory reporting requirement, though universal registration for clinical trials should be encouraged [12]. It also stipulates that, in general, it must be registered no later than 21 days after the first participant's enrollment. In addition, the International Committee of Medical Journal Editors (ICMJE) endorses the registry of clinical trials by making it mandatory for consideration for publication in most reputable journals [13]. Finally, there are civil monetary fines and the possibility of suspending and/or withholding federal research funds for compliance failure. Aside from the important ethical and scientific reasons for mandatory registry, the site, as it is a publicly available database, can be an important source of information for potential study subjects and referring physicians, which may facilitate subject recruitment.

Before the study begins, commit to publishing the study, irrespective of the trial's results. The Trial and Experimental Studies Transparency (TEST) Act that is before the US Congress at the time of this chapter's writing will likely strengthen the FDA's Section 801 requirements by broadening the scope of studies that must register with ClinicalTrials.gov, placing deadlines on when studies must be published following trial completion, and requiring more trial results to be reported to the registry [14, 15].

If the study is sponsored by industry, particular attention must be paid to the collaborative research agreement that is signed between the industry sponsor and the institutions and investigators carrying out the research. The conditions of the agreement must be ethical and agreeable to all parties involved, particularly regarding data and editorial control. Individual pharmaceutical companies often have guidelines on collaboration with academia in clinical trials freely available on the web and ought to be reviewed prior to entering any sort of agreement, as well as details of the individual agreement [16, 17]. Most importantly, ensure that the investigators, and not the sponsors, retain control over the data and the content of the manuscript. The ICMJE has taken a strong stand on this point [18]. In addition, while most agreements require that the sponsor be able to review the content of the study, a limit to the amount of time that they have to do so may prevent intentional and unintentional delays in submitting the manuscript for publication in the case of disappointing results.

Authorship can be a challenging subject to discuss, but the conversation is made much easier when it is done prior to beginning the trial. Authorship, including the efforts necessary to achieve authorship, should be agreed upon as early as possible. To be considered an author, an investigator must have meaningful involvement in

the trial design, conduct, data management and analysis, writing, and editing. Specifically, authorship should be "based on (1) substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; (2) drafting the article or revising it critically for important intellectual content; and (3) final approval of the version to be published. Authors should meet conditions 1, 2, and 3" [19]. Inappropriate authorship, be it honorary or ghost authorship, is not consistent with sound research ethics and should not happen. Finally, all coinvestigators must be up-front about conflicts of interest, financial or otherwise, in a timely and transparent manner. While definitions and reporting requirements vary across journals, most reputable journals have specific policies on conflict of interest available for public review [20].

Prior to beginning a study of human subjects, the investigators must obtain approval from their institutional review board (IRB). For multicenter trials, approval from each sites' IRB is required, and approval from one is insufficient to cover the other sites. Whether or not informed consent is required from study subjects depends on the nature of the study. Waiver for consent in clinical trials is most commonly seen in the setting of emergency research and quality improvement initiatives; however, the IRB rather than the individual investigator ultimately makes that decision. Documentation of IRB approval or waiver for informed consent is necessary. Increasing scrutiny is being placed on investigators on this topic and medical journals are more consistently requiring this documentation [21–23].

9.3 Conduct of the Trial

In the conduct of a clinical trial, the study team must maintain meticulous documentation of events and results throughout the course of the trial. Transparency with all institutional, state, and federal regulatory agencies and reporting adverse events and protocol deviations in a timely and open manner will prevent damaging allegations of misconduct and prevent the trial from being shut down prematurely. The conduct of the trial must proceed exactly as outlined in the study protocol and approved IRB application. All changes to the protocol must be IRB-approved, and substantive changes should be noted in the manuscript that is submitted for publication.

All clinical trials are required to have monitoring for safety. The nature of that monitoring "should be commensurate with risks" to the participants [24]. A formal Data and Safety Monitoring Board is generally needed for any controlled trial that evaluates mortality or major morbidity as an endpoint, is a multicenter trial, is intended to provide definitive information regarding the safety and efficacy of an intervention and/or if there are serious concern for study subject safety, whether it is the nature of the study population, the intervention being studied, or both [25]. The DSMB should be comprised of experts, including a statistician, a clinician with knowledge relevant to the condition being evaluated in the trial, and an expert in the medical specialty associated with the major potential adverse events of the intervention. The DSMB's main tasks are to evaluate periodically the results of the trial for

decisions on efficacy and to evaluate continuously for safety. The DSMB must be independent of those running the trial and, in the case of industry sponsorship, independent of the sponsor. The independence requirement and documentation to that effect are increasingly expected to be reported to the journals [26]. In addition, the monitoring plan and statistical adjustments to prevent inflation of type I error should be determined prior to starting the trial.

9.4 Analysis

The primary endpoints that are analyzed and presented must be consistent with what was registered on ClinicalTrials.gov. In two separate studies, nearly one-third of trials that had adequately registered the intended primary outcomes resulted in a publication that presented a different or discrepant primary outcome, most often one with a statistically significant result [27, 28]. Any post hoc analyses should be clearly described as such. As with multiple subgroup analyses and post hoc analyses, a lack of consistency with reporting primary outcomes determined a priori introduces a significant risk of type I error, often in favor of the intervention being tested when, in fact, the significance is completely random [29].

It is good practice to have an independent statistician reviewing the statistical software code, the statistical approach, and the interpretation to ensure that all are appropriate and accurate. In addition, many journals will require statistical review by their own in-house statistician prior to publication.

9.5 Writing

One of the goals of publication is to allow for reproducibility of the findings; accordingly, a clear explanation of what was done during the conduct of the trial is critical. In addition, clear and concise writing can make the difference between acceptance in a high impact factor journal and a lower impact factor journal. Specific guidelines, such as the Consolidated Standards of Reporting Trials (CONSORT) 25-item checklist (Fig. 9.1), have been developed that can be helpful during the writing process for randomized clinical trials to ensure clarity and reproducibility [30]. In addition, the ICMJE has detailed guidance on the content of manuscripts for clinical trials [31].

In general, the manuscript should focus on what is novel about the study and what it adds to the body of medical literature. Write clearly and succinctly, with minimal jargon, and keep in mind the journal's target audience. For this reason, many investigators choose the journal to which they will be submitting their manuscript prior to beginning the writing. When choosing a journal, it may be helpful to review back issues to get a sense of whether or not the trial is consistent with the topics covered by the journal. Asking more senior colleagues for recommendations or contacting journal editors directly can be helpful in gauging a journal's interest

Section/Topic	Item No	Checklist item	Reported on page No
Title and abstract			
	1a	Identification as a randomi sed trial in the title	
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	
Introduction			
Background and objectives	2a	Scientific background and explanation of rationale	
	2b	Specific objectives or hypotheses	
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	
Participants	4a	Eligibility criteria for participants	
	4b	Settings and locations where the data were collected	
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they wereactually administered	
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	
	6b	Any changes to trial outcomes after the trial commenced, with reasons	
Sample size	7a	How sample size was determined	
	7b	When applicable, explanation of any interim analyses and stopping guidelines	
Randomisation:			
Sequence generation	8a	Method used to generate the random allocation sequence	
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	
Blinding	11a	If done, who was blinded after assignment to interventions (for example,participants, care providers,those	
CONSORT 2010 checklist			
		assessing outcomes) and how	
	11b	If relevant, description of the similarity of interventions	
Statisticalmethods	12a	Statistical methods used to compare groups for primary and secondary outcomes	
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	
Results			
Participant flow(a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	
	13b	For each group, losses and exclusions after randomisation, together with reasons	
Recruitment	14a	Dates defining the periods of recruitment and follow-up	
	14b	Why the trial ended or was stopped	
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	
Discussion			
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	
Other information			
Registration	23	Registration number and name of trial registry	
Protocol	24	Where the full trial protocol can be accessed, if available	
Funding	25	Sources of funding and other support (such as supply of drugs) role of funders	

Fig. 9.1 CONSORT 2010 checklist of information to include when reporting a randomized trial

in and likelihood of publishing the trial results. Every major journal has their author instructions freely available on their prospective website. Follow those instructions regarding length and format exactly as they are stated. Failure to do so may result in rejection of the manuscript without further review.

For those whose primary language is not English, it may be necessary to utilize a professional editing service or to enlist the help of coauthors or colleagues with a

strong command of the English language. Be sure that there is no patient-identifying information in the manuscript. Many journals offer online supplements/appendices to the printed article. Take advantage of this chance to ensure transparency and reproducibility of the research.

The title should be descriptive of the study design and the scientific question being asked. If the trial is a randomized controlled trial, then that should be stated in the title. Avoid catchy titles and puns—these are best reserved for other settings, such as editorials. Typically, the rest of the manuscript will contain the abstract, introduction, methods, results, discussion, and funding/disclosures section.

Often it is best to write the abstract after completing the rest of the manuscript. It should begin with one or two sentences describing the context and importance of the trial, followed by a general overview of the methods of the trial. The results component of the abstract should specifically state the findings of the prespecified primary outcome. The conclusion should offer interpretation of the findings and their significance. The overall direction of the manuscript, from introduction to the discussion sections, should parallel what is written in the abstract.

The introduction of the manuscript should not be a thorough literature review, nor should it describe what is already well known by the readership. Rather, it should adequately provide the context and frame the motivation for undertaking the trial. The hypothesis should be stated clearly and unambiguously.

The methods section should be explicit about the trial design, the primary and secondary outcomes, and the arms of the study. Clearly defining the inclusion and exclusion criteria allows the reader to draw appropriate conclusions regarding the study's external validity. The randomization process, sample size calculations, and statistical analysis methods should be delineated in the methods section. To describe the study simply as "randomized" without further information is inadequate reporting. The details of the randomization process are important for the reviewers and readers to determine the possibility of bias, particularly selection bias. Whether patients were placed in a study arm by random assignment, through randomized blocks, or an adaptive allocation processes must be described, as should any stratification strategies that were employed. Allocation concealment, if any, from patients and or study staff should be described as well. Measures taken to reduce bias in the randomization and allocation concealment process strengthen a study's internal validity, and surprisingly, many authors fail to report them even when they were performed [32].

The results should begin by accounting for all patients in the study, most often in the context of a flow diagram. Baseline demographics of the subjects in the various study arms should be described. The results of the analyses that were prespecified, including the primary and secondary outcomes, are to be presented in this section. Any other analyses should be presented here as well, but their post hoc status should be clearly stated. Any deviations from the trial protocol and/or usual practice should also be explained.

The discussion section should begin with a brief summary of the findings of the study, followed by a comparison of the findings with those of other previously

published studies. Conclusions drawn about the study must be consistent with the results of the trial. Avoid the tendency to over-extrapolate the study's implications. An honest appraisal of the trial's limitations and strengths should be delineated in this section as well.

Images, figures, and tables represent an opportunity to express the components and findings of the study in a visual manner. Data are often more easily presented visually than in text and are remembered better by the reader. In fact, many readers (as well as reviewers and editors) focus on these and the abstract before or in place of reading the text, underscoring the importance of making them as visually compelling as possible. Numerous texts and articles have been published on the effective presentation of data, and the reader is referred to them for in-depth details [33–36]. Figure 9.2 demonstrates examples of good and bad visual displays of data. In general, the figures should be clear, as simple as possible to convey effectively the intended message, be of high resolution, and be submitted in the target journal's preferred format. Some journals will allow color printing, but may charge the authors for that amenity. Figures ought to be stand-alone, meaning that the reader should be able to understand the investigator's message without having to read the rest of the article.

References are most easily handled with the reference management software of the author's choice. There are many options on the market of varying capabilities and price, and most allow the user to choose the reference style to be consistent with the target journal's requirements, avoiding the hassle of manually reformatting them.

During the writing process, the help of a trusted colleague who is not intimately involved in the trial can be invaluable. For those immersed in the details of a trial, some aspects may seem clear or obvious, but that may not be true for those who have not read the protocol and been involved in the trial. A colleague's feedback about the level of detail provided in the manuscript and the clarity of the writing can help point out areas worthy of a more careful explanation. It goes without saying that a thorough review for grammar and spelling, ideally by more than one person, is mandatory prior to submitting the manuscript. While the authors will have a chance to review proofs prior to publication, inattention to spelling and other grammatical details may raise questions in the reviewer's minds as to the investigators' attention to details in the planning, conduct, and analysis of the trial as well.

While most people recognize that copying text verbatim from someone else's work without utilizing quotation marks constitutes plagiarism, fewer recognize that plagiarism is the use of another's words or ideas without appropriate attribution or permission. There are various forms of plagiarism, including plagiarism of ideas; plagiarism of text verbatim; mosaic plagiarism, which involves weaving one's own words and ideas into those of another without proper citation; and self-plagiarism [37]. Perhaps under-recognized by researchers, utilizing one's own previous work without appropriate citation is considered plagiarism. The availability of software today allows journals to easily and accurately identify plagiarism. At best, it can result in a delay in publication or manuscript rejection, though it constitutes

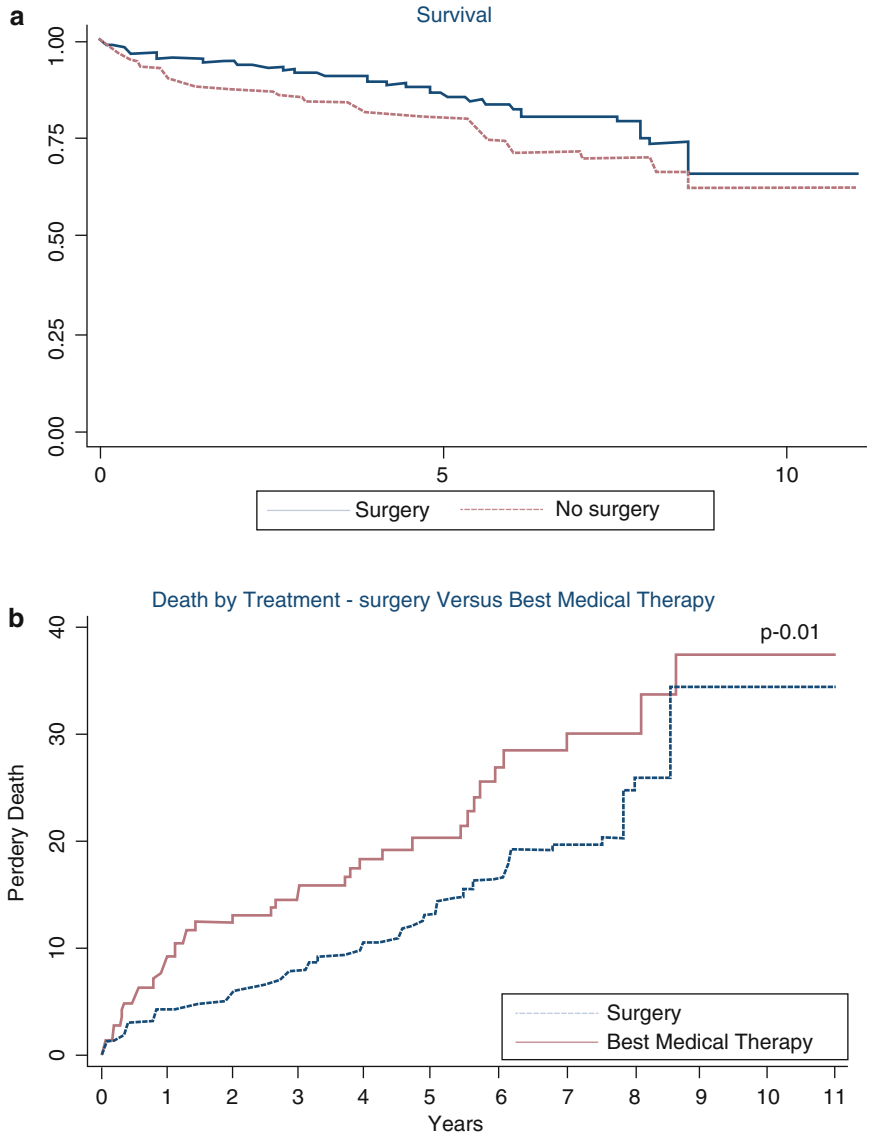


Fig. 9.2 Poor and good examples of visual displays of fictitious data. Panel **a** displays survival data, but the figure is mostly empty space. The title is not descriptive, and the legend and the axes lack captions. Panel **b** displays the same data; however, instead of displaying survival, death data are shown, which, in this case, allows for better use of space. In addition, the legend and title are more descriptive. The axes have captions, and the P value for the difference between curves is displayed. In pane **c**, there are abbreviations used that are not described in the figure. The figure lacks a title, there is unnecessary use of three-dimensional effects, and the data are displayed in a way that does not allow the reader to determine the incidences of the complications. The use of color may not be helpful when printed in *black* and *white*. Panel **d** has a title, avoids unexplained abbreviations, will be helpful for the reader when printed in *black* and *white*, and allows the reader to determine the individual complication rates and still compare between repair type

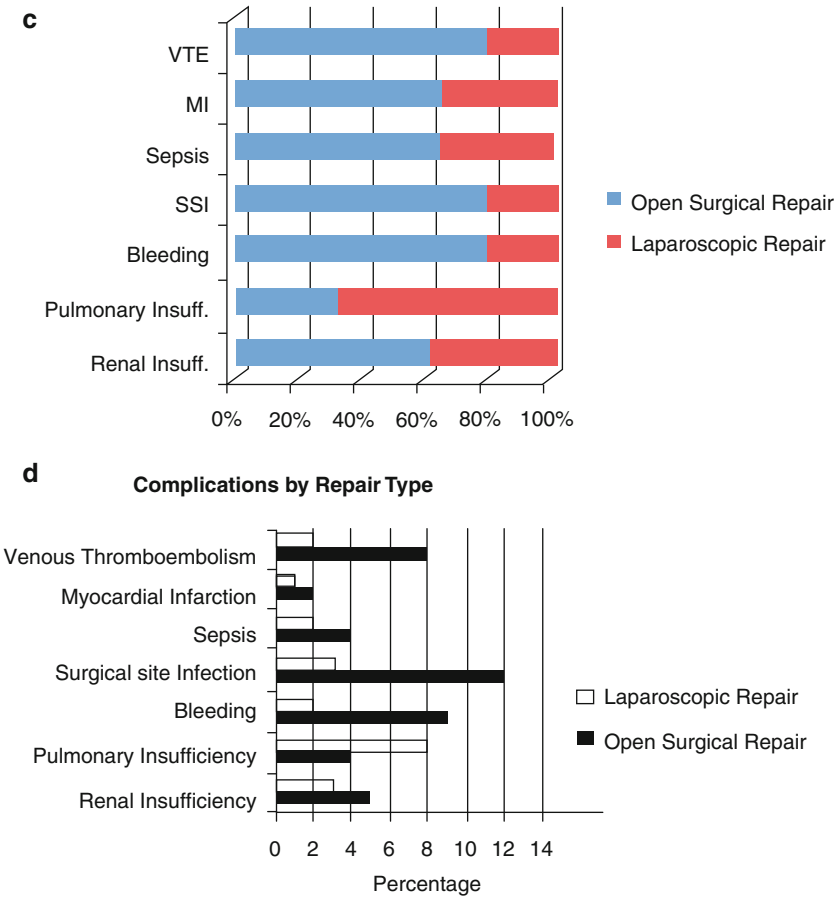


Fig. 9.2 (continued)

academic misconduct and can have even more dramatic ramifications for the offender. A thorough understanding of plagiarism is key to preventing it.

9.6 Responding to Reviewers

Very rarely is a manuscript accepted as is without revisions. Successful publications often will have required at least one round of revisions. Receiving criticism from the reviewers can be frustrating; however, the feedback should be viewed as an opportunity to improve the manuscript. Keep in mind that reviewers take the time out of their busy schedules to offer this feedback, and they are not compensated to do so. Be appreciative of their time and efforts. When responding, be sure to respond to every single point thoroughly and diplomatically.

9.7 Commit to Publishing and Do So in a Timely Manner

Regardless of the results of the trial, commit to publishing them. For previously enumerated reasons, there are ethical grounds for publishing the results of human subjects research, particularly clinical trials. Several studies have confirmed that only approximately half of all clinical trials are eventually published several years after completing the trial [38, 39]. Trials with null results are even less likely to be published, and for those that are, the delay is even longer than for trials with positive findings. The onus is on investigators to publish the results in a timely manner, irrespective of what they are.

9.8 Making the Dataset Publicly Available

There is an increasing push by federal funding agencies and medical journals for investigators to make their datasets publicly available following the conclusion of the trial and the publication of its results. If that is to be done, the data must be cleansed of all potential identifying information [40]. The following is a non-exhaustive list of potential patient identifiers that may need to be removed before a dataset can be made available to others:

- Name, including initials
- Birth date
- Any date associated with a patient (other than the year)
- Address, including partial or full zip code
- Email address
- Telephone/fax number
- Social security number
- Medical record number
- Account number
- Medical device identification/serial numbers
- Identifying photograph
- Biometric data
- Certificate/license numbers

Particular care should be taken for rare diseases or any other unusual circumstances that might make a patient identifiable.

9.9 Conclusion

The publication of a clinical trial is the culmination of tremendous intellectual and administrative efforts. While the planning, conduct, and analysis of the trial are associated with significant time and cost burdens, the publication process can be no

less challenging, though improved patient care, contribution to medical knowledge, and enhanced career success are rewards that make it well worth the effort.

References

1. Smith TL, et al. Design and results of phase I cancer clinical trials: three-year experience at M.D. Anderson Cancer Center. J Clin Oncol. 1996;14(1):287–95.
2. Estey E, et al. Therapeutic response in phase I trials of antineoplastic agents. Cancer Treat Rep. 1986;70(9):1105–15.
3. Dickersin K. The existence of publication bias and risk factors for its occurrence. JAMA. 1990;263(10):1385–9.
4. Easterbrook PJ, et al. Publication bias in clinical research. Lancet. 1991;337(8746):867–72.
5. Dickersin K, et al. Publication bias and clinical trials. Control Clin Trials. 1987;8(4):343–53.
6. Callaham ML, et al. Positive-outcome bias and other limitations in the outcome of research abstracts submitted to a scientific meeting. JAMA. 1998;280(3):254–7.
7. Krzyzanowska MK, Pintilie M, Tannock IF. Factors associated with failure to publish large randomized trials presented at an oncology meeting. JAMA. 2003;290(4):495–501.
8. Dickersin K, Rennie D. Registering clinical trials. JAMA. 2003;290(4):516–23.
9. Thornton A, Lee P. Publication bias in meta-analysis: its causes and consequences. J Clin Epidemiol. 2000;53(2):207–16.
10. Savitz DA. Failure to publish results of epidemiologic studies is unethical. Epidemiology. 2000;11(3):361–3.
11. Clinical Trials.gov. 2011. Available from: http://clinicaltrials.gov/. Cited 3 Oct 2012.
12. Congress t.U.S. Food and Drug Administration Act of 2007. Congress t.U.S., editor 2007.
13. Clinical trial registration: a statement from the International Committee of Medical Journal editors. Available from: http://www.icmje.org/clin_trial.pdf. Cited 3 Oct 2012.
14. Congress t.U.S. Trial and Experimental Studies Transparency Act of 2012. 2012.
15. Drazen JM. Transparency for clinical trials–the TEST Act. N Engl J Med. 2012;367(9):863–4.
16. GlaxoSmithKline. Public disclosure of clinical research. 2012. Available from: http://www.gsk.com/policies/GSK-on-disclosure-of-clinical-trial-information.pdf. 3 Oct 2012.
17. Merck. Merck guidelines for publication of clinical trials and related works. 2013. Available from: http://www.merck.com/research/discovery-and-development/clinical-development/Merck-Guidelines-for-Publication-of-Clinical-Trials-and-Related-Works.pdf. Accessed 11 October 2013.
18. Davidoff F, et al. Sponsorship, authorship, and accountability. JAMA. 2001;286(10):1232–4.
19. I.C.o.M.J., editors. Uniform Requirements for Manuscripts Submitted to Biomedical Journals: Ethical Considerations in the Conduct and Reporting of Research: Authorship and Contributorship. 2009. Available from: http://www.icmje.org/ethical_1author.html. 3 Oct 2012.
20. Blum JA, et al. Requirements and definitions in conflict of interest policies of medical journals. JAMA. 2009;302(20):2230–4.
21. Amdur RJ, Biddle C. Institutional review board approval and publication of human research results. JAMA. 1997;277(11):909–14.
22. Ruiz-Canela M, et al. Informed consent and approval by institutional review boards in published reports on clinical trials. N Engl J Med. 1999;340(14):1114–5.
23. Schroter S, et al. Reporting ethics committee approval and patient consent by study design in five general medical journals. J Med Ethics. 2006;32(12):718–23.
24. N.I.o. Health, editor. NIH policy for data and safety monitoring. 1998. Available from: http://grants.nih.gov/grants/guide/notice-files/not98-084.html. Accessed 11 October 2013.
25. C.f.B.E.a.R. Food and Drug Administration, Center for Drug Evaluation and Research, Center for Devices and Radiological Health, editors. Guidance for clinical trial sponsors. Establishment and operation of clinical trial data monitoring. 2006. Available from: http://www.fda.gov/downloads/Regulatoryinformation/Guidances/ucm127073.pdf. Accessed 11 October 2013.

26. Drazen JM, Wood AJ. Don't mess with the DSMB. N Engl J Med. 2010;363(5):477–8.
27. Mathieu S, et al. Comparison of registered and published primary outcomes in randomized controlled trials. JAMA. 2009;302(9):977–84.
28. Al-Marzouki S, et al. Selective reporting in clinical trials: analysis of trial protocols accepted by The Lancet. Lancet. 2008;372(9634):201.
29. Wang R, et al. Statistics in medicine–reporting of subgroup analyses in clinical trials. N Engl J Med. 2007;357(21):2189–94.
30. Trials C.S.o.R. CONSORT 2010 checklist of information to include when reporting a randomised trial. 2010. Available from: www.consort-statement.org/index.aspx?o=2965.
31. I.C.o.M.J., editor. Uniform Requirements for Manuscripts Submitted to Biomedical Journals: Manuscript Preparation and Submission: Preparing a Manuscript for Submission to a Biomedical Journal. 2009. Available from: http://www.icmje.org/manuscript_1prepare.html. 4 Oct 2012.
32. Devereaux PJ, et al. An observational study found that authors of randomized controlled trials frequently use concealment of randomization and blinding, despite the failure to report these methods. J Clin Epidemiol. 2004;57(12):1232–6.
33. Pocock SJ, Travison TG, Wruck LM. How to interpret figures in reports of clinical trials. BMJ. 2008;336(7654):1166–9.
34. Wainer H. How to display data badly. Am Stat. 1984;38:137–47.
35. Kelleher C, Wagener T. Ten guidelines for effective data visualization in scientific publications. Environ Model Software. 2011;26:822–7.
36. Freeman JV, Walters S, Campbell MJ. How to display data. How to series. 1st ed. Wiley, John & Sons, Oxford, UK, Incorporated; 2008.
37. Das N, Panjabi M. Plagiarism: why is it such a big issue for medical writers? Perspect Clin Res. 2011;2(2):67–71.
38. Hopewell S, et al. Time to publication for results of clinical trials. Cochrane Database Syst Rev. 2007;2:MR000011.
39. Ross JS, et al. Publication of NIH funded trials registered in ClinicalTrials.gov: cross sectional analysis. BMJ. 2012;344:d7292.
40. Hrynaszkiewicz I, et al. Preparing raw clinical data for publication: guidance for journal editors, authors, and peer reviewers. BMJ. 2010;340:c181.

Chapter 10
Device Versus Drug Clinical Trials: Similarities and Important Differences

T. Clark Gamblin and William S. Ragalie

10.1 Introduction

The approval processes in the United States for new drugs and devices intended to treat a medical condition share similarities; however, they routinely take different regulatory routes. The approval process mandated by the US Food and Drug Administration is a substantial part of this process. The academic investigator should have a general understanding of the steps required to bring a new drug or device to market. In an era where academic investigators often work closely in collaboration with industry in the development of drugs or devices, a working knowledge and fluency with the required regulatory steps will facilitate a successful team and provide valuable medical insights to the process.

In the United States, the authority to grant market approval to drugs and devices is vested in the Food and Drug Administration (FDA). Formed in 1906, the FDA is an agency of the US Department of Health and Human Services of the federal government. Approval of a new drug or device is based on the results of clinical investigation, often randomized controlled trials (RCT). RCTs are a relatively recent tool, being developed over the last century. In the 1930s concern over researchers' bias led statisticians R.A. Fisher and A.B. Hill to write texts on ideal design of clinical trials. The passage of the Food, Drug, and Cosmetic Act in 1938 enabled the FDA to grant approval based on demonstration of safety and efficacy through RCTs.

T.C. Gamblin, MD, MS (✉) • W.S. Ragalie, MD
Department of Surgery, Medical College of Wisconsin, Milwaukee, WI, USA
e-mail: tcgamblin@mcw.edu; wragalie@mcw.edu

T.M. Pawlik, J.A. Sosa (eds.), *Success in Academic Surgery: Clinical Trials*,
Success in Academic Surgery, DOI 10.1007/978-1-4471-4679-7_10,

10.1.1 Overview of Randomized Controlled Trials

A randomized controlled trial proceeds through several defined steps, described below.

Phase 0—A drug is given to a small number of human subjects to determine whether it behaves in humans as expected from preclinical studies. Gathering safety and efficacy data is not a goal of a phase 0 trial. Phase 0 trials are not necessary prior to beginning a phase 1 trial, but they have the ability to rapidly develop a drug by quickly establishing its behavior in humans.

Phase 1—A drug is tested in a small number of volunteer subjects, who are usually healthy, though some phase 1 trials are run with diseased patients. The purpose of phase 1 is to gather data on safety as well as dosing. Note that devices need not pass through a phase 1 trial because it would be unethical to implant devices in healthy subjects, and such a trial would not provide clinically useful information.

Phase 2—A drug or device is tested on a larger group of subjects with the disease in question. This is compared against the standard of care. Further data on safety as well as data on efficacy of the intervention are gathered.

Phase 3—This phase of investigation involves a very large number of subjects, from several hundred to several thousand. Data continue to be gathered on safety and efficacy, as well as minor adverse effects. Investigators use results of phase 3 trials in applying for market approval of the drug or device.

Phase 4—Postmarketing surveillance. This phase gathers data on safety and minor adverse events as the newly approved product is consumed in the market.

10.2 Steps Specific to the Clinical Trial of a Drug

10.2.1 Step 1: Investigational New Device (IND) Application

Clinical trials for drugs are evaluated by the largest of the FDA's five divisions, the Center for Drug Evaluation and Research (CDER). Once an investigator has obtained initial data from laboratory studies, an Investigational New Drug (IND) application is filed with the FDA. The investigator may begin the proposed clinical trial 30 days after submission of the IND. This 30 day period permits the FDA an interval in which they may mandate a clinical hold if they deem it necessary. There are two main legal ramifications of an approved IND application. It allows the investigator an exemption to the federal law requiring that a drug must be market-approved before it can be transported across state lines. In addition, it allows the drug to legally be tested in humans.

There are two categories of INDs. A "research" or investigator-initiated IND is submitted by an individual physician or investigator. An IND is categorized as "commercial" when it is submitted by a corporate entity or by one of the institutes of the National Institutes of Health (NIH).

An IND can be considered as one of three types when it is filed:

Investigational IND. This is the most commonly filed IND by academic investigators when the investigator initiates, conducts, and supervises the administration of the investigational drug.

Emergency IND. This filing allows the FDA to approve the investigational drug in cases where normal review timelines would be too long for the patients in question. This includes an after-hours emergency contact line and may be used in cases for compassionate care.

Treatment IND. This filing is used by investigators for drugs showing initial promise in clinical trials for the treatment of immediately life-threatening conditions before a final FDA review.

The CDER provides resources and guidance documents to help the researcher prior to IND submission. These include the Pre-Investigational New Drug Application Consultation Program. An index of resources provided by the FDA and their respective contact information are included at the end of the chapter.

10.2.1.1 Information Required of the IND

Certain information in three critical areas must be addressed by the investigator in the IND application:

Toxicology and pharmacology studies. The investigator's preclinical laboratory data must demonstrate the safety of the drug and explicitly state that the experimental drug will not expose trial subjects to unnecessary risk. If the drug has been approved for use outside the United States, data from other countries can be included as part of the application.

Investigator information and clinical protocols. The investigator is required to state the intention to adhere to IND regulations, institutional review board (IRB) oversight, and to obtain informed consent from all study subjects. The investigator also states the qualifications of the physicians overseeing the drug's administration.

Manufacturing information. Basic information about the physical process of creating the drug is provided. An attestation to the ability to manufacture a consistent product must describe the internal quality control, as well as ability to provide adequate quantities of the drug.

10.2.2 Step 2: New Drug Application (NDA)

The approved IND application allows the investigator to conduct and complete the proposed trial. More lengthy trials will need to include a predetermined interim analysis as well as to assess safety and efficacy. The data gathered in the

investigational phase is submitted as part of the New Drug Application (NDA) to the FDA. Provided the FDA grants approval of a NDA application, commercial use is then pursued. The NDA must contain the following information:

- Brief overview of the investigational phase
- Main ingredients of the drug
- Results of animal studies and pharmacokinetic information on the drug's mechanism of action, metabolites, and side effects
- How the drug is manufactured and packaged and quality assurance measures used in the process
- The proposed labeling for the packaged drug
- The safety and efficacy of the drug for its proposed use(s), and how the benefits outweigh the risks

The FDA provides detailed guidance on the preparation of NDAs online: www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm121568.htm

10.3 Steps Specific to the Clinical Trial of a Device

Whereas drugs are approved through the CDER, device approval takes place through the FDA's Center for Devices and Radiological Health (CDRH). This arm of the FDA regulates those who produce, repackage, relabel, or import medical devices for market use in the United States.

10.3.1 Device Classification: Classes I, II, and III

Medical devices are classified into three categories which have varied regulatory demands.

10.3.1.1 Class I

The least regulatory control is required for these devices. They are NOT intended to support or sustain life, prevent impairment, and must not present a potential unreasonable risk of illness or injury. Class I devices are subject only to general controls. Examples include simple hand held surgical instruments. General controls as defined by the FDA include registration, medical device listing, manufacturing in accordance with good manufacturing practices as defined by Title 21, Part 820 of the Code of Federal Regulations (21 CFR 820), labeling, and submission of premarket approval 510(k).

10.3.1.2 Class II

The safety of these devices cannot be assured with general controls alone and is also subject to special controls as defined by the FDA, which include specific labeling requirements, mandatory performance standards, and postmarketing surveillance. Examples of class II devices are powered wheelchairs, x-ray systems, and surgical drapes.

10.3.1.3 Class III

These devices are subject to the greatest degree of regulatory control. The purpose of such devices is to improve or sustain human life, and their use may pose a substantial risk to the patient or others. Implantable cardiac defibrillators, nerve stimulators, and breast implants are included in this category. The CDRH has determined that general and special controls alone are insufficient to ensure safety of class III devices. These devices require the review process of premarket approval (PMA), described later.

An investigator can determine the class of an investigative device by locating a description of the device in Title 21 of the Code of Federal Regulations Parts 862–892 (21 CFR 862–892), or may search the Product Classification Database through the FDA's web page, listed in the index. This compendium contains close to 2,000 descriptions of devices grouped into 16 medical specialties. The investigative device is matched with a general description of an existing device in 21 CFR, which states the class and intended use of the device as well as marketing requirements.

10.3.2 Regulatory Control for Class II and III Devices: Investigational Device Exemption, Premarket Approval, and Premarket Notification 510(k)

An important distinction between drug and device approval is that devices do not require phase I clinical testing, because it would not be clinically useful to implant devices in subjects without the targeted disease and thus be considered unethical. Class I devices are also exempt from the additional regulatory review of premarket approval (PMA) and Premarket Notification 510(k).

In addition to complying with standard clinical trial regulations, investigators studying a class III device must submit an investigational device exemption (IDE) to allow shipment of the device prior to beginning a clinical trial. The IDE is submitted to the institutional review board (IRB) and must be approved by an institutional review board prior to beginning the clinical trial. If the device is felt to pose substantial risk, the IDE must be approved by the FDA as well. Requirements for the IDE are detailed in Title 21 of the Code of Federal Regulations, Part 812 (21 CFR 812).

10.3.3 Premarket Approval Application

The premarket approval application (PMA) is the most stringent and thorough regulatory application required for market approval of a device. Regulations governing PMA are found in 21 CFR, Part 814. The officially stated review period of a PMA is 180 days, though in practice, it may be longer. After a decision has been rendered regarding the application, the FDA issues a notice and provides interested parties 30 days to petition the decision.

The FDA states that "good science and scientific writing is a key to approval of PMA application." There are two main technical sections to the application:

10.3.3.1 Nonclinical Laboratory Studies Section

This includes laboratory and animal testing data gathered prior to enrollment of human subjects (e.g., data on immunology, microbiology, metabolites, shelf-life durability). This section demonstrates that the trial was conducted in compliance with Good Clinical Practice for Nonclinical Laboratory Studies, as defined by 21 CFR Part 58.

10.3.3.2 Clinical Investigations Section

This section must include the study protocols, data collection and analysis, determination of safety and efficacy, and adverse reactions. Individual patient data and experiences must be tabulated in this section as well.

The FDA provides guidance documents for PMA preparation that are device-specific (see Sect. 10.5).

Not all class III devices require a PMA. To determine whether this is necessary, the investigator can search a description of the device in the Product Classification Database. If the description of the device in the CFR contains the language "No effective date has been established of the requirement for premarket approval," then the investigator need not submit a PMA, but a Premarket Notification 510(k) will be required instead.

10.3.4 Premarket Notification 510(k)

A 510(k) application is a premarket submission made by the investigator to the FDA. Technically, a 510(k) is required of all class I, II, and III devices that do not require PMA. However, many class I and II devices are exempt from 510(k) submission. The goal of the 510(k) application is to demonstrate that the investigative

device is "substantially equivalent" in safety and efficacy to a device that is (1) already legally marketed and (2) not subject to PMA. The grounds on which an investigative device is judged to be substantially equivalent to an existing device are based on the devices intended use, technological characteristics, and safety and efficacy profiles. After the 510(k) has been approved by the FDA and the investigative device is deemed substantially equivalent to an existing product, the device may be legally marketed. Denials of 510(k) may be appealed by addition of new data to the application or through a reclassification petition.

FDA requirements regarding who must submit the 510(k) for an investigative device are legally complex and tied closely to the manufacturing and ownership process. The parties that must submit a 510(k) are the following:

US manufacturers or foreign manufacturers with US representatives when introducing a device to market in the United States.

Entities that do not manufacture the device but repackage or relabel the device in such a way that it significantly alters its condition (e.g., significant alteration of package insert, prepackaging sterilization techniques). Most relabeling entities are not required to submit a 510(k).

Specification developers who market the device in the US market. A specification developer is one who outsources all or most of the components of the manufacturing process to other firms and who specifies the conditions on how the device is to be manufactured.

A 510(k) may also be required when a developer makes significant changes to an already approved device. The FDA's policy regarding this matter is referenced in the index.

10.3.5 Class I and II Exemptions to 510(k)

There are many exemptions to 510(k) submission for class I and II devices. All substantially equivalent devices legally marketed before May 28, 1976 that have not been significantly altered are exempt. In addition, many class I devices are exempt from good manufacturing practices. A reference to these exempt devices is included in the index.

10.3.6 Combination Products

An investigator may attempt to bring a product to market that contains elements of both a drug and a device. In this case, the investigator files all applications through the FDA Office of Combination Products. The investigator can also file a Request for Designation (RFD) through the FDA for guidance on application requirements.

10.4 Categorical Similarities and Differences Between Drug and Device Trials

10.4.1 Training of Investigators

The success of a drug trial does not heavily depend on the technical skills of an investigator, whereas a device trial (e.g., endovascular stenting) may depend greatly on the technical skills of the administering personnel. As such, technical training of personnel should be written into the study design of a device trial, as well as standards describing how device implantation will be monitored. Investigator training in drug trials should emphasize how to instruct patients to take the drug, as well as recognizing and reporting possible adverse events.

10.4.2 Adverse Events

Requirements regarding the reporting of adverse events differ for drugs versus devices. Because of the systemic effects of drugs, any adverse event that occurs must be captured and analyzed, initially, as potentially related to the investigative drug. Although the mechanism of action is known and adverse events should be identified in preclinical testing, clinical trials will also serve to identify the full spectrum of activity of the drug. In a device trial, all adverse events must be reported, but it is likely that many will be deemed unrelated. For example, a bowel obstruction in a subject who underwent implantation of a cardiac defibrillator may not be deemed related to the device, but a myocardial rupture would be.

10.4.3 Reimbursement

Usually, drugs are provided free to investigator sites and patients, and the cost is borne by the sponsor of the clinical trial. The development cost for most devices, especially class III, can be prohibitively expensive to allow free distribution. At the same time, payers may choose not to reimburse for a device that is not yet market-approved. However, investigators are able to legally seek reimbursement for use of investigational devices. In 1995 the Congress approved Medicare reimbursement for investigational devices. Previously, Medicare did not reimburse investigational devices undergoing trial. The FDA places “experimental devices” into two categories for potential reimbursement. Please note this classification is distinct from class I, II, and III classification.

10.4.3.1 Category A: Experimental

These devices are novel technologies in which the "absolute risk" has not been established due to the device's innovative, first-of-a-kind nature.

10.4.3.2 Category B: Investigational, Nonexperimental

These devices are innovations to existing, approved devices, and the fundamental questions of safety and efficacy have been answered.

After the FDA has categorized the investigational device, it informs the Centers for Medicare and Medicaid Services (CMS), who then consider potential reimbursement for the device. This arrangement applies only to CMS and does not apply to private insurers. Many private insurers do not reimburse for investigational devices and will only consider the matter on a case-by-case basis.

The orchestration of a clinical trial is complex and challenging and requires the coordination of a full team of investigators, data analysts, and ancillary personnel. The regulatory forms are extensive and require a great deal of time. The FDA offers guidance to investigators following the steps outlined in the chapter. The index includes a variety of resources and contacts that the investigator may consider during the process. Our hope is that this information will serve as a useful reference tool for those who are interested in bringing new drugs and devices through clinical trials and to market for the benefit of the patient of tomorrow.

10.5 Index

1. Product Classification Database
   (a) http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfPCD/classification.cfm
2. Premarket Approval (PMA) Application Guidance Documents
   (a) http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/default.htm
   (b) http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/HowtoMarketYourDevice/PremarketSubmissions/PremarketApprovalPMA/ucm143067.htm
3. When to Submit a 510(k) for a Change to an Existing Device
   (a) http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm080235.htm

4. Devices Exempt from 510(k) and GMP requirements http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpcd/315.cfm
5. How to Prepare a 510(k)
   (a) http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/HowtoMarketYourDevice/PremarketSubmissions/PremarketNotification510k/ucm134572.htm
6. IDE Exemption FAQs http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/HowtoMarketYourDevice/InvestigationalDeviceExemptionIDE/ucm051480.htm
7. CDER Main Page
   (a) http://www.fda.gov/Drugs/UCM2018538
8. Investigational New Drug Application (IND) Overview
   (a) http://www.fda.gov/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/InvestigationalNewDrugINDApplication/default.htm#Laws,%20Regulations,%20Policies%20and%20Procedures
9. Overview of Device Regulation
   (a) http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/Overview/default.htm
10. IND Application Main Page
    (a) http://www.fda.gov/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/InvestigationalNewDrugINDApplication/default.htm
11. FDA Resources on Drugs and Devices
    (a) http://clinicaltrials.gov/ct2/info/fdalinks

Requirements for Device Approval

Establishment registration—21 CFR Part 807
Medical device listing—21CFR Part 807
Premarket Notification 510(k)—21 CFR Part 807 Subpart E
Premarket approval (PMA)—21 CFR Part 814
Investigational device exemption (IDE)—21CFR Part 812
Quality System Regulation (QS)/good manufacturing practices (GMP)—21 CFR Part 820
Labeling—21 CFR Part 801
Medical device reporting—21 CFR Part 803

Device Approval for Emergency Use Contacts

Working hours—301.796.3400
After hours—1.866.300.4374 or 301.796.8240

Suggested Reading

Fisher RA. The design of experiments. London, Oliver & Boyd. 1935.
Hill AB. Principles of medical statistics. London, Lancet Limited. 1937.
Blake-Michaels M. Considerations for medical device trials. Appl Clin Trials. 2010; 19(6):32–8.
Arbit HM. How to prepare an investigational device exemption (IDE) as a sponsor-investigator. SoCRA Source. 2009;62(4):58–63.
Blair Holbein ME. Understanding FDA regulatory requirements for investigational new drug applications for sponsor-investigators. J Invest Med. 2009;57(6):689–95.

Chapter 11
Defining the Study Cohort: Inclusion and Exclusion Criteria

Lisa M. McElroy and Daniela P. Ladner

11.1 Introduction

This chapter focuses on the importance of defining the study cohort by means of appropriate eligibility criteria (inclusion and exclusion) in the design of a clinical trial. In statistics, a cohort is a group of individuals with common characteristics who are initially defined and composed, then examined or tracked over a given time period [1]. The term cohort can also be used when the membership of a group is defined by a non-time-based factor, for example, gender. In a clinical trial, the study cohort is also referred to as the study group or subjects. Defining the study cohort begins with clearly defining study-specific eligibility criteria. The National Institutes of Health (NIH) defines eligibility criteria as "the standards that determine whether individuals should be permitted to enter a clinical study," encompassing both inclusion and exclusion criteria [2]. It is essential that these criteria are well-defined and appropriate to answer the key questions of the study. Eligibility criteria should (1) be clear, such that eligibility of subjects can be determined easily; (2) be practical, allowing for feasible recruitment of the required sample size; (3) permit the study results to be generalizable to the population; and (4) establish the ethical foundation of the study. Study subject characteristics to consider when defining inclusion and exclusion criteria might include age, gender, comorbidities, treatment history, and place of residence.

L.M. McElroy, MD • D.P. Ladner, MD, MPH (✉)
Division of Transplantation, Department of Surgery, Northwestern University Transplant Outcomes Research Collaborative (NUTORC), 676 N. St. Clair Street, Suite 1900, Chicago, IL 60611, USA

Department of Surgery, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA
e-mail: lisa.mcelroy@northwestern.edu; dladner@nmh.org

T.M. Pawlik, J.A. Sosa (eds.), *Success in Academic Surgery: Clinical Trials*, Success in Academic Surgery, DOI 10.1007/978-1-4471-4679-7_11,

11.2 Inclusion Criteria

Inclusion criteria are the list of requirements that all study subjects have to meet in order to qualify for a clinical trial. Inclusion criteria determine which study subjects are required to meet the study goal of the clinical trial and are defined during the design phase of the clinical trial. To optimize the ability to attribute the measured study outcomes to the tested intervention, inclusion criteria need to be chosen consciously to minimize the impact of confounding characteristics or variables on the outcome of the clinical trial. Study subjects included in the study cohort of a clinical trial should be representative of the general population or target population for the intervention and be able to develop the outcome of interest. Therefore, characteristics of study subjects need to be appropriately matched with the central goal of the study in mind. For example, inclusion criteria may consider age, level of fitness, menstrual cycle phase, use of specific medications, risks to develop certain disease states, and tobacco use.

In addition, study subjects need to be accessible for enrollment into the clinical trial. Furthermore, study subjects need to fully understand the nature of the study intervention and have the ability to provide informed consent and most importantly be willing to participate in the clinical trial.

11.3 Exclusion Criteria

Exclusion criteria outline which individuals should not be enrolled into the clinical trial, regardless of their potential to develop the outcome of interest. Exclusion criteria serve to protect individuals who are at high risk of developing adverse effects from the study intervention, as well as to minimize confounding of the study outcomes by individuals with excessive medical comorbidities. Exclusion criteria must be chosen so as to maintain the balance between defining a study population that is best suited to answer the study question and maintaining the largest pool possible of individuals that are eligible for enrollment. Poorly written or vague exclusion criteria expose the study to bias; therefore, all exclusion criteria must be clearly and strongly justified with specific rationale [3]. Exclusion criteria should be listed as positive statements, such as “prior diagnosis of hypertension,” rather than negatively stated inclusion criteria such as “no history of hypertension.” Individuals who fail to adhere to pretest requirements, suffer from extensive comorbidities that complicate attributing study outcome to the intervention, or are unlikely to be available for follow-up should be excluded from the study cohort. In addition, individuals who have been unresponsive to prior interventions or have participated in other clinical trials may compromise internal validity and thus may need to be excluded from the study cohort. Vulnerable populations, such as children, pregnant women, and the elderly, or others in whom the interventions might be harmful should also be considered for exclusion, depending of the study objective [4].

11.4 Establishing Eligibility Criteria

11.4.1 Defining the Outcome

The precise definition of the study outcome of interest is an essential aspect of clinical trials, as it guides the formulation of the inclusion and exclusion criteria of the clinical trial. For example, the measurement of hypertension as an outcome of interest necessitates the strict definition of hypertension for the study. The threshold of the systolic blood pressure in one instance might be defined as above 140 and above 80 mmHg for the diastolic blood pressures, while in another instance it might be defined as above 150 and above 70 mmHg. Furthermore, the systolic and diastolic threshold values might be considered independently in some studies, while other studies may require both the systolic and diastolic blood pressure threshold to be exceeded to meet the definition of hypertension as an outcome. Alternatively, hypertension can be defined relative to a set of baseline measurements in study subjects, such as the increase of systolic blood pressure greater than 10 mmHg above the baseline values of the study subject. In addition to strictly defining study outcomes, the means by which the outcomes are measured need to be clearly outlined using the analogy of hypertension: Who will provide the blood pressure measurements (primary care provider, specialist)? Where will the measurements be performed (outpatient care hospitalizations)? How will the blood pressure measurements occur (invasive monitoring, blood pressure cuff)? Since blood pressure measurements vary over time, it must be defined over which period of time and how many successive measurements need to occur for a potential subjects to meet inclusion criteria. In which subjects can accurate blood pressure measurements be obtained and recorded? Factors such as body mass index, the availability of blood pressure measurement equipment, measurement technique, and the cohort baseline prevalence of the measured outcome (hypertension) all need to be taken into consideration when delineating the recruitment of the study subjects.

11.4.2 Study Sample Size

The size of the study cohort that is required to detect a measurable difference between intervention and control group depends on the nature of the measured condition, how precise of an intervention effect is desired, the availability of the study participants, and the ability to follow up with the study participants over the desired length of time. The power calculation determines the ability the study sample size has to detect an actual difference of the measured outcome between the control and intervention group (study arm). Preliminary data from prior research can serve to estimate event rates (e.g., hypertension) in control and intervention groups. The treatment effect, or effect of the intervention, is the difference between the event rate in the control and intervention study arms. Generally for the power calculation,

a difference is considered factual when the risk of detecting an effect where none exists is less than 5 % ($p<0.05$) and the power to detect such a difference ranges between 80 and 90 %. These parameters are used to calculate the size of study cohort that is needed to carry out the study and detect an effect of the intervention, if such an effect exists. Clinical trials that aim to detect a large treatment effect often require very large sample sizes. It is therefore advisable, in an effort to maintain recruitment feasibility, to determine the smallest cohort size required to demonstrate a significant treatment effect. The determination of the sample size has to take into account logistical challenges and differences between the control and experimental group, such as loss to follow up, drop out, or nonadherence of study subjects to the intervention [5, 6]. Clinical trials that require large sample sizes must employ effective techniques and strategies to attract and retain participants, which may include educational sessions about the clinical trial, videos or interactive computer programs conveying the importance of the study question, as well as financial incentives [7]. Using continuous rather than categorical variables to measure outcomes can reduce the sample size required. For example, if the outcome of interest is hypertension, the recording of the actual blood pressures (continuous variables) rather than the binary hypertension yes/no increases the power to detect changes to the blood pressures due to the intervention and therefore allows using a smaller sample size. Other effective approaches to reduce the required sample sized for clinical trials may involve paired measurements, where study subjects act as their own controls (time control), or the recruitment of additional study subjects into the control group. Finally, preliminary findings from a clinical trial can be further explored with a clinical follow-up study that employs larger sample sizes and therefore detects smaller differences between the control and intervention group [8].

11.4.3 Ensuring External and Internal Validity

Internal validity is provided when the measured changes between the intervention and control groups can be ascribed to the study intervention. External validity is provided when the measured changes due to the intervention apply to and can be reproduced in the general population. While internal validity is dependent on the homogeneity of the study cohort, external validity depends on heterogeneity. The inclusion and exclusion criteria together determine how heterogeneous the study sample is, and thus the internal and external validity of the clinical trial results. The narrow use of exclusion criteria leads to homogeneity, which can improve the internal validity but may compromise external validity or the generalizability of the study to the general population. While broad inclusion criteria increases external validity and may facilitate the recruitment of study subjects, this may also introduce inconsistency into the study cohort and thereby increase the likelihood of confounding of the study results. Thus the inclusion and exclusion criteria must be chosen to achieve a balance between ensuring the accuracy of the study results and the generalizability of those results to the population at large. Small pilot clinical trials

benefit from cohort homogeneity where treatment differences can be more easily demonstrated, while larger clinical trials require sufficient heterogeneity to prove generalizability [9].

11.4.4 Maintaining Feasibility of a Clinical Trial

To ensure the study feasibility, study and protocol design as well as inclusion and exclusion criteria must be clearly delineated, keeping study location, recruitment method, and individual patient factors in mind. To maintain the feasibility of a clinical trial, the following questions should be considered: What level of recruitment support may be required to offset protocol design challenges? What kind of investigators is most likely to provide high recruitment for the clinical trial? Does the study design deter study subjects or specific groups (e.g., Hispanics, women) from participation, and how might this be changed or mitigated? Can the study afford to prioritize certain groups of study subjects over another? What are the projected enrollment rates for various study sites and recruitment methods? Once the study inclusion and exclusion criteria have been established, the study protocol needs to be optimized in terms of study sites, recruitment methods, and access to potential study subjects. The time, expertise, and resources required to successfully recruit study participants are frequently underestimated and can lead to delays or disruptions in study if recruitment is not optimally planned for [10].

11.4.5 Study Cohort Recruitment

There is an array of recruitment methods that are employed for optimal recruitment of subjects into human research studies. For large clinical studies with broad selection criteria, common recruitment methods include the use of advertisements, such as newspaper, radio, and television ads. Also telephone reminders, monetary incentives, and providing additional study information have proven effective. Some research organizations maintain a database of potential participants, where consent is provided ahead of time by potential subjects, allowing them to be contacted for research studies. Smaller studies with more narrow selection criteria may employ more directed methods such as approaching an investigators own patients, students, or employees via a third party, or performing a medical record review to identify prospective subjects who will then be contacted and asked to participate in the study either in person or by telephone or mail. Large-scale epidemiological studies and other population-based studies may identify study subjects through registries, multi-institutional medical record review, or national databases. In order to ensure the feasibility of the clinical trial, recruitment methods, the cost, and access to potential study subjects need to be carefully considered when defining the study selection criteria [11, 12].

11.5 Reporting Selection Criteria

When reporting the results of a clinical trial, it is imperative that the investigators denote which study subjects were studied and how they were selected. Well-defined, consistent selection criteria allow for ease in reporting. It also allows the reader of consecutive manuscripts to understand which population the effects of the clinical trial intervention apply to and, thereby, which of their patients might benefit most from intervention. Unfortunately, many clinical trials inadequately characterize the study population or poorly justify their exclusion criteria, thereby significantly reducing their value [3, 13]. The Consolidated Standards of Reporting Trials (CONSORT) is an international group of investigators which assembled for the first time in 1993 to develop initiatives for improved reporting of randomized controlled trials. The CONSORT statement, which includes a 25-item checklist and flow diagram (Fig. 11.1), is an evidence-based set of recommendations for the standardized reporting of results from RCTs in a complete and transparent fashion that assists the critical appraisal and interpretation of the study results [14].

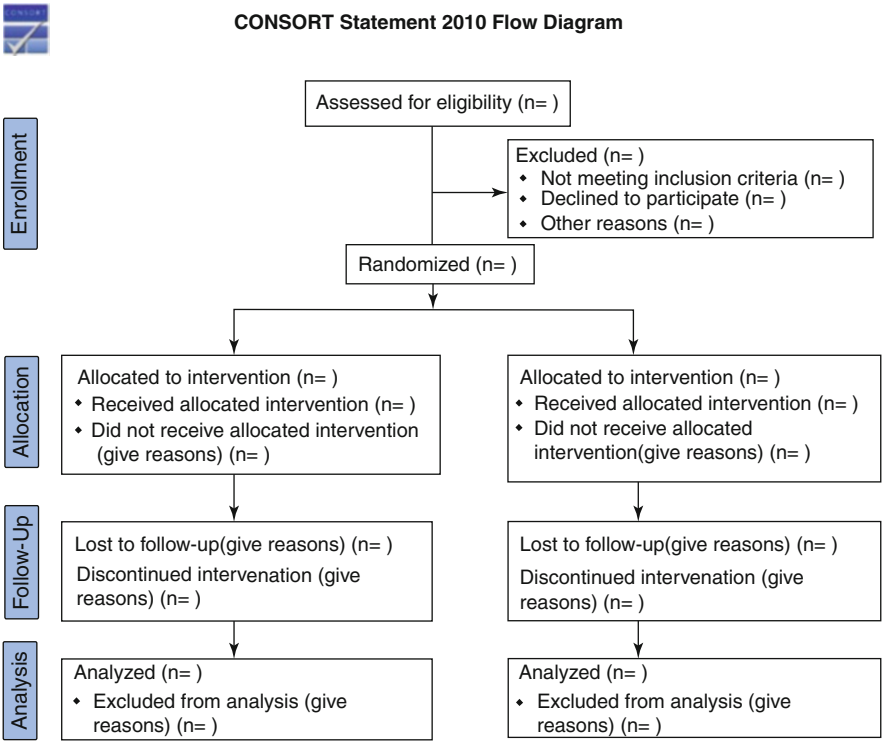


Fig. 11.1 CONSORT statement 2010 flow diagram

11.6 Modifying Selection Criteria

Under certain circumstances, it may be necessary to modify the original study eligibility criteria after recruitment has been initiated. Although the post hoc modification of study eligibility criteria may increase recruitment, such alterations can result in challenges of interpretation of the study results, as there might be significant differences in the study cohort prior and after the change in the study protocol. For this reason, the modification of study selection criteria should be regarded only as a last resort, as it carries the risk of compromising the integrity and the safety of the clinical trial [4].

11.7 Ethical Considerations

In addition to being appropriately tailored to the individual study, selection criteria must meet certain baseline ethical criteria. The safety of the participants must be considered with respect to both their baseline level of health and the possibility of experiencing adverse events as part of their participation in the study. In addition, all subjects must have the capacity to understand the nature of the study in order to provide truly informed consent. Several national standards are in place to guide researchers on the ethical inclusion of special populations in research.

11.7.1 The Belmont Report

The Belmont Report, issued by the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research in 1977, protects vulnerable populations from systematic inclusion in research and also protects vulnerable groups such as women and minorities from systematic exclusion from research. This report provides the basic ethical principles and guidelines for the conduct of research with human subjects, including clarification about the distinctions between medical practice and research. The Belmont Report put forth three basic ethical principles: first, respect for persons (individuals should be treated as autonomous agents, and persons with diminished autonomy are entitled to protection); second, beneficence (persons are to be treated in an ethical manner not only by respecting their decisions and protecting them from harm but also by making efforts to secure their well-being); and third, justice (who should receive the benefits of research and bear its burdens?). The principles of the Belmont Report have been incorporated into every aspect of human research and serve as the basis for ethical regulations in clinical trials today [15].

11.7.2 Women and Minorities

The NIH Guidelines on the Inclusion of Women and Minorities as Subjects in Clinical Research encourage the inclusion of women and racial and ethnic minorities as subjects in clinical trials. The intent of the NIH guidelines is to ensure that both the burden and the benefits of clinical trials are evenly distributed throughout society and require that all NIH-funded clinical trials determine the effect of the study intervention on both men and women and study subjects from diverse racial and ethnic backgrounds. Research plans must therefore address (1) the targeted/planned distribution of the study subjects by sex/gender and racial/ethnic groups, (2) the selection criteria of the study subjects and the rationale for the selection of sex/gender and racial/ethnic study subjects for the proposed study design in relation to the scientific objectives, (3) a compelling rationale if the exclusion of any sex/gender or racial/ethnic group is proposed, and (4) a description of the proposed outreach programs for recruiting study subjects of both sex/gender and racial/ethnic groups. In some situations, it may be acceptable to exclude study subjects based on gender or race. Examples of acceptable justifications include (1) inappropriate burden to the participants' health, (2) research questions that are only relevant to one sex/gender or racial/ethnic group, (3) if sufficient data already exist for one sex/gender or race/ethnicity, and (4) if preliminary evidence strongly suggests no difference between sex/gender and racial/ethnic groups [16].

11.7.3 Inclusion of Children

For research studies, a child is defined as an individual under the age of 21 years by the NIH. The NIH Policy on Inclusion of Children requires that research protocols provide either a plan to include children in the proposed clinical trial or, if children are excluded from the clinical trial, to present a justification for the exclusion. If children are included in the clinical trial, the study plan must include the rationale for selecting children and the selected age ranges. The study plan must also describe the expertise the study team provides to manage children of those ages and the suitability of the study facilities for children. The sample size of the recruited children needs to be large enough to contribute meaningful results for the study. The NIH guidelines state that it is expected that children be included in all research involving human subjects; several exclusionary circumstances apply: (1) the study question is not relevant to children, (2) there are laws or regulations which prohibit the inclusion of children in research, (3) the knowledge being sought is already available for children or is being obtained from another ongoing study, and (4) a separate, age-specific study in children is warranted and preferable. Other reasons to exclude children from recruitment into a clinical trial include insufficient available data to estimate the potential risks of the intervention for children or if the study design is aimed at collecting further data on pre-enrolled adult study subjects [17].

References

1. Szklo M. Population-Based Cohort Studies. Epidemiol Rev 20;1998:81–90.
2. Learn about clinical studies. At http://clinicaltrials.gov/ct2/info/understand#types. Accessed 1 Mar 2013.
3. Van Spall HG, Toren A, Kiss A, Fowler RA. Eligibility criteria of randomized controlled trials published in high impact general medical journals: a systematic sampling review. JAMA. 2007;297:1233–40.
4. Chin R, Lee BY. Patient selection and sampling: key components of clinical trials and programs. In: Principles and practice of clinical trial medicine. Boston: Academic; 2008.
5. Schulz KF, Grimes DA. Sample size calculations in randomised trials: mandatory and mystical. Lancet. 2005;365(9467):1348–53.
6. Kirby A, et al. Determining the sample size in a clinical trial. Med J Aust. 2002;177:256–7.
7. Caldwell PHY, Hamilton S, Tan A, Craig JC. Strategies for increasing recruitment to randomised controlled trials: systematic review. PLoS Med. 2010;7(11):e1000368. doi:10.1371/journal.pmed.1000368.
8. Browner WS, Newman TB, Cummings SR, Hulley SB. Estimating sample size and power. In: Hulley SB, Cummings SR, Browner WS, Grady D, Hearst N, Newman TB, editors. Designing clinical research: an epidemiologic approach. 3rd ed. Philadelphia: Lippincott Williams & Wilkins; 2001. p. 65–84.
9. Nottage M, Siu L. Principles of clinical trial design. J Clin Oncol. 2002;20:42s–6.
10. Stoney CM, Johnson LL. Development and conduct of studies. In: Principles and practice of clinical research. Amsterdam: Elsevier; 2012.
11. Mapstone J, Elbourne D, Roberts IG. Strategies to improve recruitment to research studies. Cochrane Database Syst Rev. 2007;(2):MR000013.
12. Treweek S, Mitchell E, Pitkethly M, Cook J, Kjeldstrom M, Johansen M, Taskila TK, Sullivan F, Wilson S, Jackson C, Jones R, Lockhart P. Strategies to improve recruitment to randomised controlled trials. Cochrane Database Syst Rev. 2010;(4):MR000013.
13. Friedman LM, Furberg C, DeMets DL. Study populations. In: Friedman LM, et al., editors. Fundamentals of clinical trials. New York: Springer Science and Business Media/LLC; 2010.
14. Moher D, Schulz KF, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. JAMA. 2001;285: 1987–91.
15. The Belmont Report. Office of the Secretary, Ethical Principles and Guidelines for the Protection of Human Subjects of Research; The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. 1979. At http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.html. Accessed 1 Mar 2013.
16. NIH Policy and Guidelines on the Inclusion of Women and Minorities as Subjects in Clinical Research – Amended. 2001. At http://grants.nih.gov/grants/funding/women_min/guidelines_amended_10_2001.htm. Accessed 1 Mar 2013.
17. NIH Policy and Guidelines on the Inclusion of Children as Participants in Research Involving Human Subjects. 1998. At http://grants.nih.gov/grants/guide/notice-files/not98-024.html. Accessed 1 Mar 2013.

Chapter 12
The History of Clinical Trials

Sanziana Roman

12.1 Introduction

Clinical trials are sets of tests in medical research aimed at determining efficacy and safety for various health interventions, including drugs, diagnostics, surgical procedures, devices, or clinical protocols. The reason why clinical trials have become standard in modern medicine may have been best articulated by Mark Twain: "It is best to prove things by actual experiment; then you know; whereas if you depend on guessing and supposing and conjectures, you never get educated." [1]

12.2 Early History

Clinical trials have evolved over millennia. One of the earliest documented experiments is presented in the Book of Daniel (verses 11–16, cca 600 BC, in Babylonian times):

> Then said Daniel to Melzar, whom the prince of the eunuchs had set over Daniel, Hananiah, Mishael, and Azaria: Prove thy servants, I beseech thee, ten days; and let them give us pulse to eat, and water to drink. Then let our countenances be looked upon before thee, and the countenance of the children that eat of the portion of the king's meat: and as thou seest, deal with thy servants. So he consented to them in this matter, and proved them ten days. And at the end of ten days their countenances appeared fairer and fatter in flesh than all the children which did eat the portion of the king's meat. Thus Melzar took away the portion of their meat, and the wine that they should drink; and gave them pulse. [2]

S. Roman, MD, FACS
Department of Surgery, Duke University School of Medicine,
DUMC #2945, Durham, NC 27710, USA
e-mail: sanziana.roman@duke.edu

T.M. Pawlik, J.A. Sosa (eds.), *Success in Academic Surgery: Clinical Trials*,
Success in Academic Surgery, DOI 10.1007/978-1-4471-4679-7_12,

In many ancient cultures, medicine relied on the occult and the religious. One of the classic examples of this was the belief that epilepsy was a divine intervention and was thought to be a "sacred" disease. The form of medicine that arose in fifth-century Greece, associated with the name of Hippocrates, and later popularized by Galen in Rome, marked a major innovation in the treatment of disease. Unlike supernatural theories of disease, Hippocrates' method involved seeking the causes of illness in natural factors. This method rested upon an analogy between the order of the universe and the composition of the body's "humors." Health was a matter of achieving equilibrium between competing humoral forces. Although Hippocratic theory would later be challenged, it persists today in various traditions of holistic medicine [3].

An important body of literature which influenced medicine and treatment of disease were a collection of textbooks, lectures, research, and philosophical essays on various medical subjects written by numerous authors over many decades, which ranged in time from the last decades of the fifth century BC and the first half of the fourth century AD. They are collectively known as the *Hippocratic Corpus* [4] (Fig. 12.1). A significant proportion of the writings are made up of case histories. One of the important aspects of the *Hippocratic Corpus*

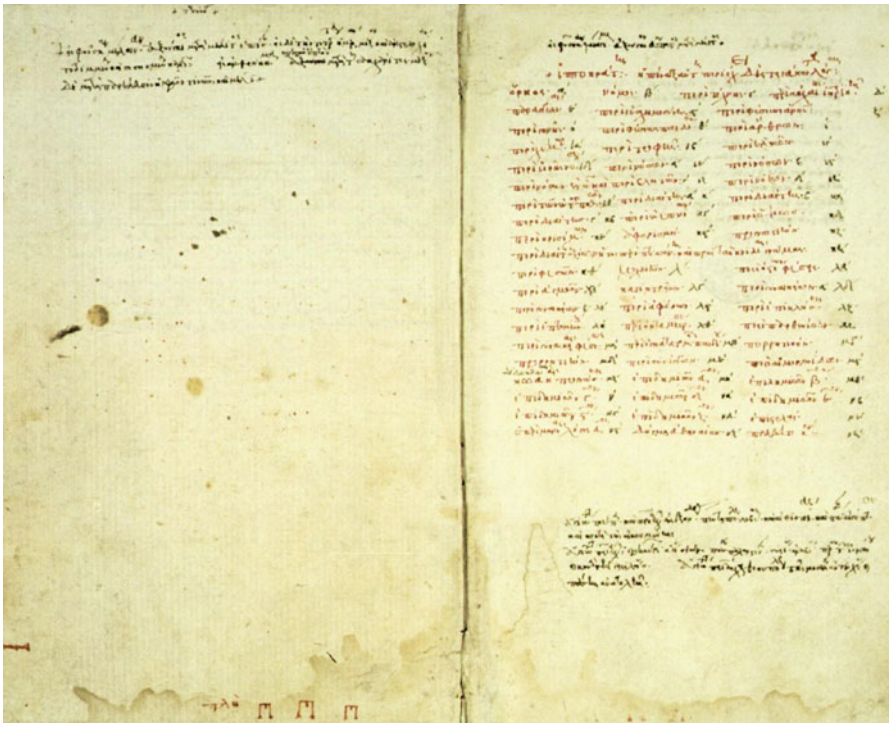


Fig. 12.1 Vaticanus graecus 277, 10v-11r: table of contents in a fourteenth-century Hippocratic Corpus manuscript. Marcus Fabius Calvus owned this manuscript, transcribed it in his own hand, and used it in the preparation of his 1,525 Latin translation (*Source*: Online Vatican exhibit [30])

is that it proved that physicians could reflect on their observations and actions. The most famous work in the *Hippocratic Corpus* is the Hippocratic Oath, a landmark declaration of medical ethics. The Hippocratic Oath not only deals with abstract principles but also practical matters, such as aiding one's teacher financially [5]. It is quite complex and likely reflects the compilation of several authors. It is most famous for the maxim which has guided physicians over millennia and, in many ways, marks the basis of all clinical trials: "As to diseases, make a habit of two things—to help", [clinical effectiveness] "or at least to do no harm" [good safety profile].

In 1025 AD, the Persian physician Avicenna (Abū 'Alī al-Ḥusayn ibn 'Abd Allāh ibn Sīnā) wrote the widely used medical treatise "The Canon of Medicine" in which he laid down rules for the experimental use and testing of drugs. He wrote a precise guide for practical experimentation in the process of discovering and proving the effectiveness of medical drugs and substances [6]. Some of these experimental recommendations included the time of action must be observed, so that effect and accident are not confused; the effect of the drug must be seen to occur constantly or in many cases, and if this did not happen, it must have denoted an accidental effect; and the experimentation must be done with the human body, for testing a drug on a lion or a horse might not prove anything about its effect on man. "The Canon of Medicine" was a very popular treatise and was used extensively in medical schools across Europe as late as 1650 [6].

12.3 Clinical Trials Begin

The earliest and most well-described prospective clinical trial was by James Lind in 1753 in his "A Treatise on The Scurvy." His methods and findings are described in great detail:

> On the 20th May, 1747, I took twelve patients in the scurvy on board the Salisbury at sea. Their cases were as similar as I could have them. They all in general had putrid gums, the spots and lassitude, with weakness of their knees. They lay together in one place, being a proper apartment for the sick in the fore-hold; and had one diet in common to all, viz., water gruel sweetened with sugar in the morning; fresh mutton broth often times for dinner; at other times puddings, boiled biscuit with sugar etc.; and for supper barley, raisins, rice and currants, sago and wine, or the like. Two of these were ordered each a quart of cyder a day. Two others took twenty five guts of elixir vitriol three times a day upon an empty stomach, using a gargle strongly acidulated with it for their mouths. … Two others had each two oranges and one lemon given them every day. These they eat with greediness at different times upon an empty stomach. They continued but six days under this course, having consumed the quantity that could be spared.…
>
> The consequence was that the most sudden and visible good effects were perceived from the use of the oranges and lemons; one of those who had taken them being at the end of six days fit for duty. The spots were not indeed at that time quite off his body, nor his gums sound; but without any other medicine than a gargarism or elixir of vitriol he became quite healthy before we came into Plymouth, which was on the 16th June. The other was the best recovered of any in his condition, and being now deemed pretty well, was appointed nurse to the rest of the sick. [7]

Fig. 12.2 Edward Jenner vaccinating James Phipps, May 14, 1796: lithograph by French artist Gaston Melingue (1840–1914) (From Horne [31])

One of the significant contributors to medicine and clinical trials was Edward Jenner (1749–1823), a British physician who was the first to prove that inoculation with cowpox could prevent deadly smallpox. He observed that milkmaids, who often would contract cowpox and developed sores on their hands, did not contract smallpox. On 14 May 1796, Jenner tested his hypothesis by inoculating an 8-year-old boy, James Phipps, who was the son of his gardener, with purulent material scraped from the cowpox blisters on the hands of his milkmaid (Fig. 12.2). Phipps developed a fever but no major illness. Later, he injected Phipps with smallpox

material several times, but the boy remained well. He repeated this experiment on 23 additional subjects and proved that immunization with cowpox could prevent smallpox. Based on these experiments, the British government banned "variolation" in 1840 and supported the use of cowpox as a widespread vaccine, free of charge [8].

The concept of randomization, blinding, and placebo controls was introduced by Amberson in 1931. He used a coin flip to determine whether two comparable groups of patients with pulmonary tuberculosis received sanocrysin or distilled water [9]. To reduce observer bias, the researchers ensured that the group assignment of the patients was known only to two of the authors of the report and the nurse in charge of the ward. Their attempts to blind the identity of the groups to which patients had been allocated using injections of distilled water were unlikely to have been successful, however, because all of the patients receiving sanocrysin suffered adverse systemic effects of the drug, including a death from liver necrosis. Amberson et al. were able to follow 19 of their 24 patients for up to 3 years after the last dose of sanocrysin, and they found no evidence of beneficial effects.

Following the report by Amberson et al. and in the same issue of the *American Review of Tuberculosis*, there was a less detailed report, in which Brock, of the Waverly Hills Sanatorium in Kentucky, arrived at very different conclusions about the effects of sanocrysin [10]. Brock concluded that the drug had "an outstanding clinical effect on exudative tuberculosis in white patients," although "very little effect in limiting the progression of the disease in black patients." These conclusions were based on his observations of 46 patients – all treated with varying doses of sanocrysin. Although the patients in Brock's study suffered some of the same toxic effects of the drug, the patients and the drug regimens with which they were treated differed from those in the study by Amberson and his colleagues. Brock's patients were not followed after the end of treatment. The stage of disease at which Brock's white and black patients started treatment differed also, as did the care of black and white patients overall, since they were treated in a segregated 1920s Kentucky Sanatorium. Prior to 1954, the 17 southern states and the District of Columbia enforced racial segregation in every area of public activity, including hospital services. Clinicians across the country recognized that Amberson's study was a better study overall, and they believed his findings. This led to the rightful demise of sanocrysin treatment for tuberculosis in the United States.

12.4 The Beginning of Large-Scale Trials

The first multicenter trials involved treatment of pulmonary tuberculosis with streptomycin and were published in the United Kingdom in 1948 and the United States in 1952. The British study encompassed 107 patients from 7 centers. The patients were carefully selected and divided into two groups: one group treated with streptomycin and bed rest, the other group with bed rest alone. They followed patients for 6 months and concluded that streptomycin-treated patients fared much better than

the control group [11]. In the United States, the Veterans Administration, together with the US Armed Services, continued multicenter trials for tuberculosis over the next two decades, with good success.

Large-scale clinical trials were viewed as becoming the gold standard for proving effectiveness of treatment. The Salk poliomyelitis vaccine trials, sponsored by the National Foundation for Infantile Paralysis (March of Dimes), started in 1954 and involved nearly 1.8 million children [12]. The trials began at the Franklin Sherman Elementary School in McLean, Virginia. Children in the United States, Canada, and Finland participated in these trials, which used for the first time the now-standard double-blind method. On April 12, 1955, researchers announced the vaccine was safe and effective, and it quickly became a standard part of childhood immunizations in America. Nonetheless, the statistical design used in this encompassing experiment prompted criticism. Eighty-four test areas in 11 states used a randomized, blinded design where all participating children aged 6–9 years received injections of either vaccine or placebo and were then observed for evidence of the disease. Other test areas in 33 states used an "observed control" design, where participating children aged 7–8 years received injections of vaccine, but in the control group, no placebo was given, and children were then observed for the duration of the polio season. The use of the dual protocol illustrates both the power and the limitations of the randomized clinical trial even in the face of legitimate therapeutic claims. The placebo-controlled trials were necessary to define the Salk vaccine as the product of scientific medicine, even though it had been supported and pushed forward by a lay activist organization (March of Dimes). However, the observed control trials were essential in maintaining public support for the vaccine as "the product of lay faith and investment in science," since placebo-controlled trials often elicited negative responses from the public [12].

12.5 Ethics in Clinical Trials

The issue of ethics with respect to medical experimentation has been an ongoing concern. One of the most blatant breaches of ethics occurred in Nazi Germany during the 1930s and 1940s. Uncovering these atrocities at the Nuremberg trials helped introduce the concept of international responsibility for medical ethics. The Nuremberg Code for human experimentation was issued in 1946 to address ethical issues surrounding human subjects' protection. This was the first document to set out ethical regulations in human experimentation based on informed consent [13]. It contained ten principles related to a physician's ethical duties and made informed consent "absolutely essential."

It was revised in a declaration adopted in June 1964 in Helsinki, Finland (the Helsinki Declaration). This was a nonlegally binding instrument in international law set in ethical principles in regard to human experimentation and was developed by the World Medical Association [14]. A notable change from the Nuremberg code was the relaxation of the conditions of consent, asking doctors to obtain consent if "at all possible," and introduce the concept of a proxy consent, such as a legal guardian. This Declaration has undergone numerous revisions over the years. The

first revision in 1975 introduced the concept of oversight of research by an independent committee, which became the system of institutional review boards (IRB). In the United States, IRB regulation became official in 1981 [14].

12.6 A Test of Medical Ethics in the United States

One of the most infamous studies and a clinical trial where ethical principles were lacking was the "Tuskegee Study of Untreated Syphilis in the Negro Male" conducted from 1932 to 1972 by the US Public Health Service with the cooperation of the Tuskegee Institute [15]. Its aim was to record the natural history of untreated syphilis. It involved 600 black men from Macon County, Alabama, including nearly 400 men with late-stage syphilis and 200 healthy controls. In return for participation, the subjects were promised free physical examinations, a meal on the day of examination, and burial stipends. The subjects were not informed whether they had syphilis or not, and many were told they were being treated for "bad blood," which was a common local lay term. Some officials at the Centers for Disease Control later believed that "bad blood" was a local synonym for syphilis; however, numerous patients interviewed over time corroborated the fact that they were not aware that they may have harbored syphilis. Prior to 1946, the standard treatment for syphilis consisted of injections of arsenic and mercury. Penicillin was found to be an effective cure for syphilis in 1946; however, the subjects enrolled in the study were not offered this known effective treatment and were not told about their syphilis diagnosis.

In 1972, Jean Heller of the Associated Press broke the story of this study. After examination of the US Public Health Service, it became obvious that the study did not have a formal protocol. While the study never claimed to be about testing treatment effectiveness, the magnitude of the risks taken with the subjects involved led numerous individuals to feel that the Public Health Service had "played" with human lives. Senator John Sparkman of Alabama denounced the study as "absolutely appalling" and "a disgrace to the American concept of justice and humanity" [15].

The Tuskegee study proved to be an American tragedy. It ultimately played a key role in creating the institutions and practices that govern the use of human volunteers in US biomedical research today. However, it made the public wary of agreeing to participate in clinical studies and introduced a certain level of distrust between patients and physicians.

12.7 The Belmont Report

On July 12, 1974, the National Research Act was signed into law, creating the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. The Commission was charged with identifying the basic ethical principles that would underlie the conduct of biomedical and behavioral research involving human subjects and developing guidelines to assure that research is

conducted in accordance with those principles. The committee met monthly for a period of 4 years. After a 4-day meeting held in February 1976 at the Smithsonian Institution's Belmont Conference Center, the Belmont Report was put forth to summarize the basic ethical principles identified by the Commission in the course of its deliberations. This report stands today as a statement of basic ethical principles and guidelines that should assist in resolving the ethical problems that surround the conduct of research with human subjects, including clarifying the distinctions between medical practice and research. It put forth three basic ethical principles: (1) respect for persons (divided into two separate moral principles: to acknowledge autonomy and to protect those with diminished autonomy), (2) beneficence (including assurances that research maximizes possible benefits and minimizes possible harm), and (3) justice (declaring that the benefits and burdens of research are divided equally among individuals) [16].

The principles of the Belmont Report have been incorporated into every aspect of human research and are the basis for ethical regulations in practice today.

12.8 The Food and Drug Administration (FDA)

The history of the FDA can be traced to the latter part of the nineteenth century and the US Department of Agriculture's Division of Chemistry. Under chief chemist, Harvey Wiley, the Division began conducting research into the misbranding of food and drugs on the American market. Wiley used these findings and his alliances with diverse state regulators and national associations of physicians and pharmacists to lobby for a new federal law to set uniform standards for food and drugs. Wiley's advocacy came at a time when the public had become sensitized to public safety issues. The 1902 Biologics Control Act had been put in place after diphtheria antitoxin was collected from horses and injected into children, resulting in several deaths from tetanus contamination [17]. In the 1920s, numerous men and women died from the ingestion of Radithor, a radium-containing drink manufactured by Bailey Radium Laboratories, New Jersey, which was touted as a magical cure all [18]. These tragic events had raised public awareness and made lobbying for food and drug safety regulation popular.

In June 1906, President Theodore Roosevelt signed into law the Food and Drug Act, also known as the "Wiley Act." The Act prohibited the transport of food which had been "adulterated," with that term referring to the addition of fillers of reduced "quality or strength," coloring to conceal "damage or inferiority," formulation with additives "injurious to health," or the use of "filthy, decomposed, or putrid" substances. The act applied similar penalties to the marketing of "adulterated" drugs, in which the "standard of strength, quality, or purity" of the active ingredient was neither stated clearly on the label nor listed in the *United States Pharmacopoeia* or the *National Formulary* [19].

In 1927, the Bureau of Chemistry's regulatory powers were reorganized under a new US Department of Agriculture body, the Food, Drug, and Insecticide

organization. This name was shortened to the Food and Drug Administration shortly thereafter. By the 1930s, journalists, consumer protection organizations, and federal regulators began mounting a campaign for stronger regulatory authority by publicizing a list of harmful products which had been ruled permissible under the 1906 law, including radioactive beverages such as Radithor, the mascara Lash Lure, which caused blindness and worthless "cures" for diabetes and tuberculosis. The resulting proposed law was signed by President F. D. Roosevelt on June 24, 1938 as the new Food, Drug, and Cosmetic Act. The new law significantly increased federal regulatory authority over drugs by mandating a pre-market review of the safety of all new drugs. The law also authorized factory inspections and set new regulatory standards for foods, cosmetics, and therapeutic devices. This law, though extensively amended in subsequent years, remains the central foundation of the FDA today [20].

12.9 Clinical Trials and Statistics Evolve

A major stimulus for clinical trials in the United States arose from the Kefauver-Harris Amendment of 1962 to the US Food, Drug, and Cosmetic Act [21]. This set forth legal requirements for "adequate and well-controlled investigations" that had to be satisfied before a drug can be approved by the FDA. It required that drug advertising discloses accurate information and stopped cheap generic drugs being marketed as expensive drugs under new trade names. This amendment was a response to the thalidomide fallout, where thousands of children were born with birth defects as a result of their mothers taking the drug for morning sickness during pregnancy. The law was signed by President J. F. Kennedy in 1962. The Drug Efficacy Study Implementation began classifying all pre-1962 drugs already on the market as effective, ineffective, or needing further study [22].

In 1976, the Medical Device Amendments to the Federal Food, Drug, and Cosmetic Act extended some of the testing requirements established for drugs to medical devices as well. They established three regulatory classes based on the degree of control necessary to assure that the various types of devices are safe and effective. This classification continues to be applicable today [23].

In the 1980s, Thomas Chalmers led a movement to eliminate potential bias when physicians are permitted access to study the data during the course of a clinical trial by separating patient care from treatment evaluation functions. Chalmers was the first in the United States to create a department of biostatistics and a department of geriatrics at the Mount Sinai Hospital in New York, where he was the director. He was a strong proponent of analytic and methodological rigor in clinical trials and contributed significantly to the science of meta-analysis.

The roots of meta-analysis may be traced back to the seventeenth-century studies of astronomy, but the first collation of all conceptually identical experiments concerning a particular research topic conducted by independent researchers was noted in the 1940 book-length publication, *Extra-sensory perception after 60 years*, authored by

psychologists at Duke University and spearheaded by J. G. Pratt [24]. This encompassed a review of 145 reports on extrasensory perception experiments published from 1882 to 1939 and included an estimate of the influence of unpublished papers on the overall effect, known as the "file-drawer problem." In the 1970s, more sophisticated analytical techniques were introduced by medical statisticians Gene Glass, Frank Schmidt, John Hunter, and Thomas Chalmers. The term "meta-analysis" was coined by Gene Glass, who described it as an "analysis of analyses" [25].

12.10 Cancer Research in the United States

The creation of the National Cancer Institute (NCI) in 1937 designated the start of federally sponsored medical research in the United States, which ultimately was recognized as the National Institutes of Health (NIH) in 1981. In 1955, the NCI formed a clinical studies panel. During one of its early meetings, Sidney Farber and others introduced the concept that the study of leukemia would move forward more quickly if physicians worked together on clinical trials through a "cooperative group" mechanism, which would allow for broad collaboration among researchers from various medical institutions. This research method already had been proven successful in Veterans Administration hospitals when they studied tuberculosis [26]. The NCI created eight cooperative cancer research groups: the Eastern Cooperative Oncology Group, the Cancer and Leukemia Group B, the Gynecological Oncology Group, the Northern Central Cancer Treatment Group, the National Surgical Adjuvant Breast and Bowel Project, the Radiation Therapy Oncology Group, the Southwest Oncology Group, and the Children's Oncology Group. These groups have had significant success over the years in several oncology areas, and they continue to exist today. To name a few of the significant contributions, they produced long-term survival and cures for a majority of pediatric cancer cases; demonstrated that breast conservation surgery is often better than radical mastectomy; developed optimal adjuvant chemotherapy regimens and hormonal therapy for breast cancer; developed paclitaxel as an efficient treatment for ovarian cancer and metastatic non-small cell lung cancer; showed that interferon is a mainstay of treatment for metastatic melanoma; and showed that combination chemotherapy and radiotherapy is the most effective treatment for advanced cervical cancer [27].

12.11 Recent Times

As clinical trials became more complex, they required more regulation and administration. Clinical trials at academic centers often ran from specific medical departments. More encompassing clinical trial offices (CTO) emerged over the last two decades to consolidate administrative activities related to clinical trials, ranging from protocol development to billing compliance. Their main goal was to enhance institutional research capabilities. A review of CTOs at eight academic health

centers in 2008 revealed, however, that there was little uniformity in the structure or functions designated to the CTOs across institutions; some were gatekeepers on all budgeting and billing, others provided educational or liaison services, while some had monitoring and auditing responsibilities for compliance [28]. This review pointed to the fact that institutions still are challenged by the lack of clearly defined organizational structure for clinical trial administration. CTOs will become increasingly important as there is increased pressure on academic organizations to focus their billing and compliance activities, increase communication between researchers, consolidate education and training, decrease costs and redundancy in infrastructure, and increase visibility of trials.

Clinical trials continue to evolve in scope, design, funding, regulation, and administration. With the overall decline in federal funding for research in the United States, physicians and researchers have developed new approaches to medical research, including looking at international collaborations and harmonization of resources. This has led to the development of large national and international mega-trials. These multicenter and multinational trials are attractive for a number of reasons, including patients' recruitment to achieve the needed numbers, regulatory and ethical issues, as well as marketing strategies. Mega-trials often include thousands of patients with significant heterogeneity in demographics, clinical characteristics, presence of comorbidities, and associated therapies in the study population. Often the need for such large sample of patients denotes the fact that the effect size of the intervention or drug is expected to be quite small. One downside to applying the findings of a mega-trial to daily medical practice is that group-averaged data are transferred to individual care often with weak demographic, ethnic, and clinical associations [29]. For these reasons, the enthusiasm for the international mega-trials has waned somewhat, and researchers have started to focus their energies in more individualized and tailored patient-centered research. While it is true that there may not be a "perfect trial," clinicians and researchers will continue to collaborate and innovate.

> Coming together is a beginning. Keeping together is progress. Working together is success. (Henry Ford)

References

1. Green S, Benedetti J, Crowley J. Clinical trials in oncology. 2nd ed. Boca Raton: Chapman & Hall/CRC; 2003.
2. Newnes G. The art Bible: comprising the old and new testaments : with numerous illustrations, vol. 1. London: G. Newnes; 1796. p. 946 (Google eBook).
3. Snowden F. Classical views of disease: Hippocrates, Galen and Humoralism, Epidemics in Western Society since 1600 (Yale HIST 234). http://www.academicearth.org/lectures/classical-views-of-disease-hippocrates-galen-and-humoralism. Retrieved 23 Dec 2012.
4. Tuke JB. Hippocrates. In: Chisholm H, editor. Encyclopaedia Britannica, vol. 13. 11th ed. Cambridge, UK: Cambridge University Press; 1911. p. 517–9.
5. Garrison FH. History of medicine. Philadelphia: W.B. Saunders Company; 1966.
6. Flannery M. "Avicenna". 2007. Encyclopaedia Britannica Online. http://www.britannica.com/EBchecked/topic/45755/Avicenna. Retrieved 9 Sept 2013.

7. Lind J. A treatise of the scurvy in three parts. Containing an inquiry into the nature, causes and cure of that disease, together with a critical and chronological view of what has been published on the subject. London: A. Millar; 1753, as transcribed by Lars Bruzelius, publically available, http://www.bruzelius.info/Nautica/Medicine/Lind%281753%29.html. Retrieved 23 Dec 2012.
8. Hopkins DR. The greatest killer: smallpox in history, with a new introduction. Chicago: University of Chicago Press; 2002.
9. Amberson JB, McMahon BT, Pinner M. A clinical trial of sanocrysin in pulmonary tuberculosis. Am Rev Tuberc. 1931;24:401–35.
10. Brock BL. The sanocrysin treatment of pulmonary tuberculosis in the white and negro races. Am Rev Tuberc. 1931;24:436–45.
11. Marshall G, Crofton JW, Hurford JV, Brent RH, Ross ID, Santon Gilmour W, Nash FA, Thomas DME. Streptomycin treatment of pulmonary tuberculosis: a medical research council investigation. Br Med J. 1948;30:769–82.
12. Meldrum M. "A calculated risk": the Salk polio vaccine field trials of 1954. Br Med J. 1998;317:1233–6.
13. Vollman J, Winau R. Informed consent in human experimentation before the Nuremberg code: Nuremberg doctors trial. Br Med J. 1996;313:1445–7.
14. Rid A, Schmidt H. The 2008 Declaration of Helsinki — first among equals in research ethics? J Law Med Ethics. 2010;38(1):143–8.
15. Jones J. The Tuskegee syphilis experiment – a moral astigmatism. In: Harding S, editor. The racial economy of science: toward a democratic future. Bloomington: Indiana University Press; 1993. p. 275–86.
16. http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.html. Retrieved 1 Feb 2013.
17. http://history.nih.gov/exhibits/history/docs/page_03.html. Retrieved 10 Jan 2013.
18. "Medicine: Radium Drinks". Time magazine. April 11, 1932. http://www.time.com/time/magazine/article/0,9171,743525,00.html. Retrieved 26 Dec 2012.
19. http://www.fda.gov/AboutFDA/WhatWeDo/History/Milestones/ucm128305.htm. Retrieved 31 Dec 2012.
20. Hamowy R. Medical disasters and the growth of the FDA. Independent Policy Reports. Published by The Independent Institute, Oakland, CA; 2010.
21. http://www7.nationalacademies.org/archives/drugefficacy.html. Retrieved 31 Dec 2012.
22. Krantz Jr JC. New drugs and the Kefauver-Harris amendment. J New Drugs. 1966;6(22):77–9.
23. http://www.fda.gov/medicaldevices/productsandmedicalprocedures/deviceapprovalsandclearances/pmaapprovals/default.htm. Retrieved 31 Dec 2012.
24. Bösch H. Reanalyzing a meta-analysis on extra-sensory perception dating from 1940, the first comprehensive meta-analysis in the history of science. In: Schmidt S, editor. Proceedings of the 47th annual convention of the Parapsychological Association. Vienna: University of Vienna; 2004. p. 1–13.
25. Glass GV. Primary, secondary, and meta-analysis of research. Educ Res. 1975;5(10):3–8.
26. http://swog.org/visitors/history.asp. Retrieved 31 Dec 2012.
27. Kelahan AM, Catalano R, Marinucci D. The history, structure, and achievement of the Cancer Cooperative Groups. Managed Care Cancer. 2001:28–33.
28. Rubin E, Lazar D. Clinical trials office: what is new in research administration? Published by the Association for Academic Health Centers; 2009.
29. Volpe M, Tocci G, Pagannone E. Fewer mega-trials and more clinically oriented studies in hypertension research? The case of blocking the renin-angiotensin-aldosterone system. J Am Soc Nephrol. 2006;17(4 Suppl 2):S36–43.
30. Avicenna (al-Husain b. Abdallah Ibn Sina, d. 1037). Online Vatican exhibit. http://www.ibiblio.org/expo/vatican.exhibit/exhibit/f-medicine_bio/Medicine_2.html (Public Domain).
31. Horne CF. Great men and famous women: a series of pen and pencil sketches of the lives of more than 200 of the most prominent personages in history. New York: S. Hess; 1894 (United States public domain).

Chapter 13
Ethics of Clinical Trials

Jukes P. Namm and Peter Angelos

13.1 Ethical and Regulatory Background

The aim of clinical trials is to obtain generalizable knowledge that can be used to advance health and health care. The ethical issues in clinical trials arise when human subjects, who may not directly benefit from the research, are faced with the risk of being exploited or harmed. The current ethical framework for research is based on guidelines and laws which were a necessary response to historical abuses of participants in the name of science (Table 13.1).

13.1.1 Early Codes of Research Ethics

Albert Neisser, a professor of dermatology and venereology at the University of Breslau and the person who first identified gonococcus, published clinical trials on serum therapy in patients with syphilis in 1898. In an attempt to find a way to prevent the spread of syphilis, he injected serum from syphilis patients into patients who were admitted for other medical conditions. Most of his subjects were prostitutes who were neither informed about the experiment nor asked for consent. After an ensuing outbreak of syphilis, he claimed that it was not due to the serum, but because the patients worked as prostitutes.

J.P. Namm
Department of Surgery and MacLean Center for Clinical Medical Ethics, University of Chicago, 5841 S. Maryland Avenue, MC 6040, Chicago, IL 60637, USA
e-mail: jukes.namm@uchospitals.edu

P. Angelos (✉)
Department of Surgery and MacLean Center for Clinical Medical Ethics, University of Chicago, 5841 S. Maryland Avenue, MC 4052, Chicago, IL 60637, USA
e-mail: pangelos@surgery.bsd.uchicago.edu

T.M. Pawlik, J.A. Sosa (eds.), *Success in Academic Surgery: Clinical Trials*,
Success in Academic Surgery, DOI 10.1007/978-1-4471-4679-7_13,

Table 13.1 Guidelines on the ethics of clinical trials

Guideline	Source	Date	Reference
Nuremberg Code	Nuremberg Military Tribunal *United States v. Brandt* et al.	1947	http://www.hhs.gov/ohrp/archive/nurcode.html
Declaration of Helsinki	World Medical Association	1964	http://www.wma.net/en/30publications/10policies/b3/index.html
Belmont Report	National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research	1979	http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.html
45 CFR 46 (Common Rule)	US Department of Health and Human Services	1991	http://www.hhs.gov/ohrp/humansubjects/guidance/45cfr46.html
Good Clinical Practice: Consolidated Guidance	International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use	1996	http://www.fda.gov/downloads/Drugs/…/Guidances/ucm073122.pdf
International Ethical Guidelines for Biomedical Research Involving Human Subjects	Council for International Organizations of Medical Sciences and the World Health Organization	2002	http://www.cioms.ch/images/stories/CIOMS/guidelines/guidelines_nov_2002_blurb.htm

In 1900, the Prussian Ministry of Religion, Education and Medical Affairs set forth the Berlin Code of Ethics as a direct response to Neisser's syphilis study. It stated that all research interventions in a medical institution could only be performed by the medical director or with his or her authorization. All medical interventions other than for diagnosis, healing, and immunization were prohibited under all circumstances if "the human subject was a minor or not competent for other reasons" or if the subject had not given his or her "unambiguous consent" after a "proper explanation of the possible negative consequences" of the intervention. These requirements, in addition to the other aspects of the case, had to be "documented in the medical history" [1].

13.1.2 Nuremberg Code

After World War II, Nazi war criminals were tried before the International Military Tribunal for crimes against humanity. The subsequent Doctors' Trial which took place on December 9, 1946, charged 23 Nazi physicians and scientists with murder and torture under the guise of scientific research. The Nazi regime labeled all Jews,

gypsies, Slavs, homosexuals, and disabled persons as subhuman which did not entitle them to basic human rights. Various experiments in the concentration camps included the dissection of live infants, castration of boys and men without the use of anesthesia, and sterilization of women with an X-ray machine, as well as the effects of high-voltage electric shocks, hypothermia, pressure chambers, and euthanasia [2].

The international outrage from these heinous crimes gave birth to the Nuremberg Code, which was drafted in 1947 as a means to prevent the future abuse of human subjects. Above all, the Nuremberg Code states that participation in research must be voluntary, and research subjects must have the ability to withdraw from the study at any time. The benefits of the research to the individual or to society must outweigh the risks, and any unnecessary suffering must be avoided. Investigators must be qualified to conduct the study and prepared to stop the study should the risks become unacceptable for the participants (e.g., the possibility of death or disabling injury as a foreseeable consequence) [3].

13.1.3 Declaration of Helsinki

The World Medical Association, in response to the atrocities of the Nuremberg Trials, issued the *Ethical Principles for Medical Research Involving Human Subjects* in 1964 and since then has undergone multiple revisions (most recently in 2008). Taking its name from the city in which it was adopted, these principles became known as the Declaration of Helsinki.

This document stresses the importance of participant health and close monitoring of subjects, especially populations that include the "economically and medically disadvantaged," those who cannot give informed consent (or who may be doing so "under duress"), those who will not benefit personally from the research, and those for whom "research is combined with [medical] care." All research subjects must have access to the "best" standard of care treatment as identified in the study, especially pertaining to placebo-controlled trials, which is an attempt to address the potential conflict in the goals of the clinician (to care for the patient) and the researcher (to obtain generalizable knowledge). It also states that voluntary, fully informed consent for all research participants is imperative [4].

13.1.4 The Jewish Chronic Disease Hospital

In 1963, Chester Southam, the chief of virology at the Sloan-Kettering Cancer Institute, injected live, cultured cancer cells into 22 debilitated patients at the Jewish Chronic Disease Hospital in Brooklyn, New York. He believed that despite the patients' old age and debilitated state, their immune systems would reject the cancer cells. The patients were never fully informed of the experiment because Southam stated that they would be of "no consequence" to them and did not want to unduly

upset them by the word "cancer." Although all of the patients eventually did reject the cancer cells as Southam predicted, the lack of informed consent and disclosure of risks, although minimal, illustrates the prevailing paternalistic attitudes in medical research at that time [5].

13.1.5 Willowbrook Hepatitis Study

Willowbrook was an institution for mentally retarded children in Staten Island, New York. Saul Krugman, an infectious disease specialist, was consulted to study immunity against hepatitis which was endemic to the institution from 1956 to 1967. The study involved feeding controlled amounts of the virus, which is shed in the feces, to healthy children. Although the parents who brought their children to Willowbrook consented to the study, they were coerced into believing that there were no beds available except for research subjects. It was also explained to them that the contraction of hepatitis was "inevitable" when in fact only 30–53 % of children acquired the disease at the institution [6]. In 1966, Henry Beecher, a professor of anesthesiology at Harvard Medical School, criticized numerous clinical trials including the Willowbrook study in his landmark article *Ethics and Clinical Research* and questioned why Krugman did not place more emphasis on promoting hygiene and sanitation to decrease the risk of infection [7].

13.1.6 Tuskegee Syphilis Study

From 1932 to 1972, the Public Health Service conducted a clinical study on the treatment and natural history of syphilis. The subjects were 399 poor and mostly illiterate African American sharecroppers with syphilis in Alabama who neither gave informed consent nor were informed of their diagnosis and the risk to others through sexual contact. They were told that they had "bad blood" and were offered free medical treatment, one free meal per day, and $50 in case of death for the funeral. In 1943, when penicillin was discovered as a cure for syphilis, the subjects in the study were never offered the drug so that the investigators could further study the natural history of the disease. The study resulted in 28 deaths, 100 cases of disability, and 19 cases of congenital syphilis [6]. The story appeared on the front page of the New York Times in July 1972. The study was terminated, and the Public Health Service was forced to settle a $9 million class action lawsuit filed by the NAACP which was divided among the participants.

13.1.7 Belmont Report

The public response to the Tuskegee syphilis study prompted the federal government to establish a National Commission in 1974 for the purpose of drafting

guidelines for the ethical conduct of research involving humans. The Belmont Report was published in 1979 and named after the conference center at the Smithsonian Institute in Washington, DC, where the National Commission laid much of the groundwork. The Belmont Report explicitly noted the principles of *respect for persons*, *beneficence*, *and justice* that have become the pillars of the ethical conduct of research.

The principle of respect for persons applies to informed consent and requires that the autonomy of the individual be respected. Those who do not have decisional capacity (e.g., children or persons with intellectual disability) require the consent of an authorized third party who is able to make decisions based on the individual's best interest.

The principle of beneficence begins with the medical precept *primum non nocere*—or first do no harm. Potential harm should be avoided if possible, but always minimized, by balancing the risks and the benefits to the research participant. Adhering to this ethical standard is not only the responsibility of the investigator but also of an independent institutional review board (IRB) to determine if the protocol is justifiable.

The principle of justice requires that both the burdens and benefits of research are distributed fairly. Historically, the burdens of research were placed upon the economically disadvantaged and the vulnerable, some of whom never reaped any physical benefits from the research. Because medical research is a public good, all of society should benefit and likewise be called upon to participate equally [8].

13.1.8 Code of Federal Regulations (45 CFR 46 and 21 CFR 50, 56)

As a response to the Belmont Report, the federal government sought to enforce the regulatory requirements of research involving human subjects in the United States. These efforts were incorporated into the Code of Federal Regulations (CFR), Title 45, Part 46 (45 CFR 46) or the Common Rule, which went into effect in 1991. The Common Rule falls under the authority of the Department of Health and Human Services (DHHS) and focuses on the process of review, approval, and oversight of research involving human subjects for all institutions receiving any federal support for research. Subpart A of 45 CFR 46 outlines the responsibilities of the IRB including membership, function, review, and necessary records to approve research protocols. It also delineates the required elements of the informed consent form and the criteria for waiving informed consent. Subparts B, C, and D involve additional protections bestowed upon pregnant women, fetuses, neonates, prisoners, and children [5].

The CFR, Title 21, Parts 50 and 56 (21 CFR 50, 56) protects the rights, safety, and welfare of subjects in clinical investigations involving products regulated by the Food and Drug Administration (FDA) which include food and color additives, drugs for human use, medical devices for human use, biological products for human use, and electronic products [9].

13.1.9 International Guidelines

The Good Clinical Practice (GCP) guidelines (International Conference on Harmonization, 1996) and the International Ethical Guidelines for Biomedical Research Involving Human Subjects (Council for International Organizations of Medical Sciences and the World Health Organization, 2002), based on the Belmont Report and the Declaration of Helsinki, were adopted as an international standard for designing, conducting, recording, and the reporting of clinical trials. Specifically, they delineate that research must be scientifically sound and must be approved by ethical review committees [10]. In 2008, the FDA published the Final Rule which replaced the requirements of the Declaration of Helsinki with those of the GCP guidelines.

13.2 An Ethical Framework for Research

With the harsh lessons from history, these multiple guidelines and regulations have helped to create a framework with which to minimize the exploitation of humans for research, while still fostering the goals of improving health and health care. Each guideline by itself is not comprehensive in addressing the ethical complexities in research. However, collectively, they uphold eight ethical principles for clinical research: *collaborative partnership*, *social value*, *scientific validity*, *fair participant selection*, *favorable risk-benefit ratio*, *independent review*, *informed consent*, *and respect for participants* [11]. Some of these principles will at times conflict, but it is not sound ethical practice to simply ignore one when two or more appear to be at odds. It is the investigator's responsibility to take all of these principles into consideration and weigh the value of each one for any individual case.

According to the recommendations in the Belmont Report which are now codified in 45 CFR 46, *research* is "*a systematic investigation, including research, development, testing and evaluation, designed to develop or contribute to generalizable knowledge*." And a *human subject* is *a* "*living individual about whom an investigator obtains data through intervention or interaction with the individual*, *or identifiable private information*" [12]. Surgical innovation is an important topic that does not cleanly fit into the definition of research, yet is not considered standard of care. This topic goes beyond the scope of this chapter, but the ethical issues of informed consent and risk assessment remain paramount as surgical innovation continues to play an important role in the progress and development of surgical care [13].

13.2.1 Collaborative Partnership

Clinical research, in its role to serve a social good, should create a partnership with the community in which the research is being conducted. Although there is no legal

requirement to do so, communication between the investigator and the community helps to prevent any exploitation and can maximize the potential benefits to the community. The goals of research can be attained through respect for the community's values, context, culture, and social practices. This partnership can be formed through community advisory boards, advocacy groups, and town hall meetings. Care should be taken to ensure that both the direct and indirect benefits of the research are fairly distributed between the research team and the community [11].

13.2.2 Social Value

Clinical research has social value because it can lead to improvements in health or health care. However, research void of social value places subjects at risk for no benefit. To have social value, the research results have to be generalizable as well as implementable. Although all research cannot be applied universally, it is important to first define which group or population will benefit from the research. Secondly, it becomes necessary to determine the potential benefit of the research for the particular group or population. Thirdly, a structure should be in place to disseminate the knowledge acquired from the research to maximize its social value. Such structures include publications in peer-reviewed journals, presentations at medical conferences, letters and presentations to the community, and press releases in the media. Lastly, research should lead to strengthening the infrastructure of health care in the community [11].

13.2.3 Scientific Validity

In order to truly benefit society, research must be based on valid science. The first crucial step in achieving scientific validity is in the design and methods of the research. The research should have clear objectives, adequate sample size, unbiased and reliable outcome measures, and appropriate statistical analysis. The feasibility of the study should be considered to ensure that the study will reach accrual in a reasonable amount of time [11].

Adequate sample size and study design are too often overlooked by investigators. According to a study looking at 54 surgical trials from 2008, 28 % of the trials reported negative results despite being underpowered. Furthermore, of those that were underpowered, 47 % of the studies did not report an a priori power calculation, and 40 % reported inappropriate interpretations of the results [14]. The use of underpowered studies to implement novel therapeutic surgical interventions without any evidence that these therapies are equivalent to the standard of care not only place patients at increased risk of unnecessary harm, but they also make the risks that the research participants undertook unjustifiable due to the studies' lack of scientific or social value.

Randomized controlled trials (RCT) are accepted as the highest standard of evidence regarding the safety and efficacy of new therapies. Placebo-controlled trials are designed to blind subjects and investigators from potential bias during an RCT comparing specific therapies. The FDA requires "adequate and well-controlled" studies to demonstrate the safety and effectiveness of drugs as a condition of approving their clinical use [3]. However, the Office for Human Research Protections states: *A design involving a placebo control should not be used where there is a standard treatment that has been shown to be superior to placebo by convincing evidence* [15]. The potential risks of placebos are not always negligible and must be taken into consideration.

Charles Fried (1974) believed that it is ethically problematic for a physician to enroll one of his or her patients in a randomized trial if the physician believes that one of the treatments which is available outside of the study is superior. Prior to enrolling a patient in a trial, Fried argued that the physician should be indifferent to either treatment (i.e., unsure which treatment is better), which is now commonly described as *equipoise* [16]. Some have argued that as a clinician, it is too difficult—if not impossible—to have absolute equipoise with regard to different treatments for a patient. Therefore, others have proposed that RCTs could still be ethical if there is equipoise among the scientific community or the community at large even if an individual physician believes one treatment may be superior to the other. Attempts to address the problem of equipoise, or lack thereof, have included using interim data analysis, unbalanced randomization, and adaptive randomization which allocates more subjects to the superior arm once data emerge. However, adaptive randomization can confound the interpretation of the results due to potential differences between the groups if one arm is enrolling significantly more subjects [17].

13.2.4 Fair Participant Selection

The target group for a research study should be decided based on scientific objectives. With the lessons learned from the numerous cases of abuse in the past, social factors that are irrelevant to the research should never influence the selection of the target group. Diligence must be taken to minimize risk to the group by excluding individuals who may be at increased risk of adverse effects from the therapy being studied. Also, participants should be chosen to maximize both the social value as well as the individual benefit to the research participants [11].

The recruitment of research participants is most commonly done through a treating clinician or through advertisements. It is important that the treating physician maintain the integrity of his or her responsibility to the care of the patient before enrolling participants into clinical trials, especially when capitation fees are collected from the study sponsor. Researchers must guard against recruiting subjects through misleading advertisements that are particularly concerning for vulnerable or desperate patients who are running out of standard treatment options. It is crucial

that such patients are not misled regarding the purpose of the research study [18]. Research involving vulnerable populations such as individuals with impaired decision-making capacity, children, and prisoners must meet a higher standard to ensure that the risks to the subjects are minimized, the outcomes of the research directly benefits the population, and the research is unable to be performed on a less vulnerable population. Current challenges in fair participant selection include underrepresentation of minorities and the elderly, as well as the overrepresentation of economically disadvantaged populations [19].

13.2.5 Favorable Risk-Benefit Ratio

For each individual research participant, the net benefit of research should outweigh the risks. In the case where the individual's risks outweigh the benefits, the social value must be able to justify the increased risk-benefit ratio. In every case, the potential risks of research must be clearly defined and should include not only physical risks but also psychological, social, and economic risks to the participant or community. In the same manner, the benefits of the research to the individual and group need to be identified. However, in the case of benefits, only health-related potential benefits should be weighed. Secondary benefits such as payments or greater access to health services should not be factored in the balance as these may falsely justify higher-risk research by increasing payment or services [11].

Phase I clinical oncology trials are a complicated situation where the research evaluates the safety and toxicity of a new drug but provides few or no benefits to the individual. An area of debate arises from the issue of the *therapeutic misconception* where the subject mistakenly believes—or is led to believe—that the research study will in some way directly benefit him or her. Empirical data examining deficiencies of disclosure to the research subject is still lacking; however, it appears that most patients who enroll in Phase I clinical oncology trials do not understand the non-therapeutic nature of the trials and still hope for a therapeutic benefit [20]. In all situations, the risk-benefit ratio, if considered acceptable by the investigator, must then be scrutinized and approved by the IRB and then ultimately by the potential subject during the informed consent process.

13.2.6 Independent Review

An independent ethical review of all clinical research is necessary to minimize the inherent conflict of interests that potentially affect all investigators as well as to ensure that the public benefit is not derived from the exploitation of other individuals. The current review system has been established into law by 45 CFR 46. The IRB is made up of at least five members including scientists, clinicians, statisticians, and members of the community who are free from any conflict of interest with the

researchers and the clinical studies [11]. The Institute of Medicine issued a report in 2002 that outlined four specific conditions that should exist in all investigational reviews: *accountability*, *resources*, *ethics education*, *and transparency in process*.

The IRB is a part of every institution's system to protect human participants in research. IRBs are accountable to the volunteer research participants whose interests they represent, the institution which requires that these protections are upheld by law, as well as to the investigator to provide a fair and thorough review of the protocol. Adequate resources are important so that the IRB can effectively and efficiently perform its duties of assembling materials for review and monitoring approved protocols. Ethics education, such as a core curriculum in research and bioethics, should be offered and encouraged for investigators and IRB members prior to beginning any research activity. Many resources are available in both web-based and traditional written format. Finally, transparency in the process of research review has now become mandatory and should be accessible to the community and its members. It is recommended that 25 % of the IRB is made up of nonaffiliated lay members from the community. This helps ensure that research participants' knowledge base, understanding of the informed consent, and attitudes toward research risks and benefits are more accurately represented in the review process [21].

13.2.7 Informed Consent

The purpose of informed consent is to respect the autonomy of the individual and to ensure that he or she is not utilized merely as a means to an end. According to Kant: "The human being, however, is not a thing, hence not something that can be used merely as a means, but must in all his actions always be considered as an end in itself" [22]. Informed consent requires that the consenting individual has the capacity to understand the various risks and benefits of the proposed research and to be able to do so voluntarily without coercion. Participants should be informed of their right to withdraw from a study at any time without penalty. Verbal and written communication should be in the native language and on an appropriate level for the participants to understand. Apart from a few exceptions such as emergency research, research with minors, or those who do not have decisional capacity, only the individual who will be the research subject can give consent to participate in a research study. In the United States, a written signature is required for consent to participate in a research study.

The US Federal Regulations (i.e., 45 CFR 46) and European Community Rules (i.e., GCP guidelines) include three essential elements of a valid informed consent: *disclosure*, *comprehension*, *and voluntariness*. With full disclosure, participants must be informed of the nature and foreseeable risks of the trial along with the therapeutic benefits that may or may not be a result of enrolling. They must be aware of all appropriate alternative therapeutic options with a right to withdraw without penalty. A statement regarding the extent to which the participant's records will be kept confidential is required. Appropriate information should be given to the participant to explain who he or she should contact for any research-related adverse

event or for any other questions about the study. There are no good methods to determine the comprehension of the participants with regard to the research. However, the consent materials should be at an appropriate reading level (8th grade reading level) and translated into the appropriate language as necessary. Finally, voluntariness involves lack of coercion whether it is by the physician's influence or through monetary means. Compensation for travel or lost wages is acceptable; however, payment above and beyond those thresholds may cause some participants to accept a higher level of risk than they normally might take on otherwise and is considered unethical [23].

Vulnerable populations should be protected by specific safeguards to ensure safety, informed consent, and absence of coercion. Individuals with impaired decision-making capacity, children, and prisoners are groups identified by the National Commission whom voluntary written informed consent were not considered feasible or seemed overly protective. In balancing the social good that could arise from the research, as well as allowing these populations to have access to its potential benefits, the National Commission determined that in light of the principles of beneficence and justice, that research with a modified informed consent or consent by proxy would be permissible. However, in order to maintain the protection of these vulnerable populations, the commission described a *necessity requirement* where the research must be relevant to the vulnerable population and cannot otherwise be done in a less vulnerable population. They also stipulated that strict informed consent could be modified or waived if the research posed *minimal risk* to the subjects. Minimal risk is defined as no more than the *physical or psychological harm that is normally encountered in the daily lives*, *or in the routine medical*, *dental*, *or psychological examination of healthy persons*. In research with children, although they cannot legally give consent to enroll in a research study, once they are of the age where they may understand the reasons for and implications of the study, in conjunction with an appropriate consent from the parents, an assent from the child should be obtained. Likewise, dissent from the child should also be appropriately respected [24].

Research in the emergently ill population—emergency research—is a specific area where obtaining informed consent is not feasible. In addition to considering the previous principles, consulting the community in order to disclose the nature of the study and seek community input are requirements for emergency research [25]. In order for the requirement of informed consent to be waived in these situations, two criteria must be met: the research cannot be realistically carried out in nonemergency settings, and the research must directly address the emergency needs of the participants involved. As soon as consent or legal authorization becomes feasible, it should be obtained from the individual or a proxy if indicated [23].

13.2.8 Respect for Participants

Researchers have a responsibility to care for individuals and treat them with respect and monitor their safety throughout the entirety of a research study and sometimes,

even after the study ends. Protocols should be structured to adjust treatments in response to adverse events and for stopping the study earlier at predetermined time points if excessive harms or significant benefits are demonstrated during interim analysis. In the event of an injury or adverse event from participation in a clinical trial, compensation remains voluntary and is usually limited to remedial medical care. Procedures to ensure the protection of participant confidentiality should be in place. Participants have a right to have their information kept confidential from the public, insurance companies, and even other family members. In the case of biobanks and genetic research, it is imperative to clearly indicate whether the participants or their family members want to be contacted in the future if any actionable discoveries are made. However, even with all the safeguards in place, participants should always be informed that there can be no guarantees that absolute confidentiality can be maintained [11].

13.3 Clinical Investigator Conduct

13.3.1 Conflict of Interest

Conflict of interest has become one of the core components in the ethics of medical research. The two major areas where conflict of interest may arise in the clinical research setting are the *physician-researcher* conflicts and *financial* conflicts.

The physician-researcher conflict revisits the issue of equipoise and the fundamental tension that arises between the duties of a physician to the well-being of the patient with the duties of a researcher to scientific inquiry. The most blatant example is in the Tuskegee syphilis study where the physicians did not offer the subjects penicillin once it was discovered 29 years prior to the closing of the trial in order to further study the natural progression of the disease. The researchers likely did not intend to harm the subjects; however, their commitment to scientific discovery clouded their clinical judgment and their primary responsibility to care for the sick. There is no way to completely remove this conflict of roles from research. However, one way to prevent this conflict from harming participants is through research ethics committees such as the IRB.

Financial conflicts have become more prevalent in the past 10 years with the potential for large profits, the rise of medical device and pharmaceutical industries, and their influence in both clinical and research activities. Companies may pay clinicians for recruitment of patients into their study (i.e., *capitation*), fund various studies, sponsor an endowed chair, or provide company stocks. These conflicts set the stage for compromising patient safety in the recruitment and during the course of the trial as well as in the integrity of the data and reporting of the study [26].

The Jesse Gelsinger case in 1999 illustrated how financial conflicts of interest could compromise the safety of human subjects. Gelsinger was an 18-year-old male who died from complications of a gene therapy trial using adenovirus as a vector to

treat a metabolic disease. The lead investigator at the University of Pennsylvania allegedly held $13.5 million in the company stock which he sold a couple years after Gelsinger's death [27]. The FDA investigation reported that the protocol's inclusion criteria was violated for this trial, side effects were underreported, and the IRB had not been properly informed of significant events or changes to the study protocol [28]. Although it is impossible to prove a causal relationship between the financial interests and the decisions that were made in this case, it illustrates the potential dangers that these conflicts can inflict on the well-being of study participants.

Financial interests also affect issues such as the design of the study, the interpretation of its results, selective publication, and ghost or gift authorship [26]. Although industry-sponsored research has contributed a considerable amount to medical knowledge and social benefit, there is a tendency for industry-sponsored studies to reach industry-friendly conclusions or only publish the results of positive studies (which is not unique to industry-sponsored publications). Some sponsors have gag clauses that prevent researchers from analyzing data or publishing results without the consent of the sponsor which may unduly limit an investigator's ability to communicate scientifically valid research results. Furthermore, it has become more common for industry sponsors to hire private personnel to analyze and draft the manuscript, and then offer it to an established researcher in the field for review and subsequent authorship. This practice—termed *gift authorship*—is an attempt to provide more credibility for the study and, in return, becomes an easy publication for the academic researcher's curriculum vitae. Gift authors are not acceptable, and investigators must take appropriate caution when dealing with industry-sponsored studies [26, 29].

The Public Health Service and the National Science Foundation have introduced regulations regarding potential conflicts of interest. Disclosure to the investigator's institution is required if researchers have invested more than $10,000 or own more than 5 % in a company. The FDA also requires disclosure of any equity interests more than $50,000 and any payments to researchers greater than $25,000 unrelated to the costs of research [26]. Currently, there is no federal regulation requiring financial disclosure to the study participants. However, a survey of patients and a separate survey of clinicians indicate that disclosure of financial conflicts is the more ethically appropriate course in order to preserve the public's trust [30, 31]. No matter how much financial interest is at stake, these conflicts can undermine the integrity of medical research and, more importantly, curtail the benefits of research that are owed to society.

13.3.2 Authorship

Authors have an unstated contract with the readers of their work that demands accountability. Authorship usually comes with credit for the work published, but it also comes with the responsibility for the integrity and validity of the research.

Authorship is also important in the academic hierarchy and has a significant role in promotions and academic tenure. As a result, disputes regarding authorship are becoming more common yet there are no good solutions for resolving these conflicts.

The International Committee of Medical Journal Editors developed the following criteria: "Authorship credit should be based on substantial contribution to conception and design, or acquisition of data, or analysis and interpretation of data; drafting the article or revising it critically for important intellectual content; and final approval of the version to be published. Each author should have participated sufficiently in the work to take public responsibility for appropriate portions of the content" [32]. The best way to address this issue may be to keep written documentation of the authorship order, responsibilities, and expectations at the start of any research endeavor [29].

13.3.3 Research Misconduct

Research misconduct is defined as purposeful *fabrication*, *falsification*, *or plagiarism* in research. These elements of fraud are not new to the scientific community. These practices are not only unethical but also illegal according to the White House Office of Science and Technology Policy which was implemented in 2005 [33]. Although relatively uncommon, such practices are most likely underreported. The downstream consequences of research misconduct are that it can compromise the integrity of published research, destroy trust between researchers, and undermine the public support of science. Falsifying results also endanger research subjects—and even patients—if adverse events are not accurately reported or drugs are approved with erroneous data [34]. Research integrity must be an institutional priority as well as an individual commitment for every investigator. Through education, policies, and mentoring, efforts to address research misconduct are vital to the future of sound science.

References

1. Vollmann J, Winau R. Informed consent in human experimentation before the Nuremberg code. BMJ. 1996;313:1445–7.
2. Taylor T. Opening statement of the prosecution, December 9, 1946. In: Annas GJ, Grodin MA, editors. The Nazi doctors and Nuremberg Code: human rights in human experimentation. New York: Oxford University Press; 1992.
3. Department of Health and Human Services, National Institutes Health and Office for Human Research Protections. The Nuremberg Code. Available from: http://www.hhs.gov/ohrp/archive/nurcode.html. Updated 7 2005; Cited 13 Dec 2012.
4. Ashcroft RE. The Declaration of Helsinki. In: Emanuel EJ, editor. The Oxford textbook of clinical research ethics. New York: Oxford University Press; 2008. p. 141–8.

5. Porter JP, Koski G. Regulations for the protection of humans in research in the United States: the common rule. In: Emanuel EJ, editor. The Oxford textbook of clinical research ethics. New York: Oxford University Press; 2008. p. 156–67.
6. Robinson WM, Unruh BT. The hepatitis experiments at the Willowbrook State School. In: Emanuel EJ, editor. The Oxford textbook of clinical research ethics. New York: Oxford University Press; 2008. p. 80–5.
7. Beecher HK. Ethics and clinical research. N Engl J Med. 1966;274(24):1354–60.
8. Beauchamp TL. The Belmont Report. In: Emanuel EJ, editor. The Oxford textbook of clinical research ethics. New York: Oxford University Press; 2008. p. 149–55.
9. Department of Health and Human Services, National Institutes Health and Office for Human Research Protections. Food and Drug Administration. Available from: http://www.hhs.gov/ohrp/humansubjects/fda/index.html. Cited 13 Dec 2012.
10. Idanpaan-Heikkila JE, Fluss SS. International ethical guidance from the Council for International Organizations of Medical Sciences. In: Emanuel EJ, editor. The Oxford textbook of clinical research ethics. New York: Oxford University Press; 2008. p. 168–73.
11. Emanuel EJ, Wendler D, Grady C. An ethical framework for biomedical research. In: Emanuel EJ, editor. The Oxford textbook of clinical research ethics. New York: Oxford University Press; 2008. p. 123–35.
12. Department of Health and Human Services, National Institutes Health and Office for Human Research Protections. The Common Rule, Title 45 (Public Welfare), Code of Federal Regulations, Part 46 (Protection of Human Subjects). Available from: http://www.hhs.gov/ohrp/humansubjects/guidance/45cfr46.html. Updated 15 Jan 2010; Cited 13 Dec 2012.
13. Angelos P. The ethical challenges of surgical innovation for patient care. Lancet. 2010;376(9746):1046–7.
14. Brody BA, Ashton CM, Liu D, Xiong Y, Yao X, Wray NP. Are surgical trials with negative results being interpreted correctly? J Am Coll Surg. 2013;216(1):158–66.
15. Department of Health and Human Services, Office Human Research Protections. Protecting Human Research Subjects: Institutional Review Board Guidebook. Available from: http://www.hhs.gov/ohrp/archive/irb/irb_guidebook.htm. Updated 1993; Cited 13 Dec 2012.
16. Fried C. Medical experimentation: personal integrity and social policy. New York: American Elsevier; 1974.
17. Joffe S, Truog RD. Equipoise and randomization. In: Emanuel EJ, editor. The Oxford textbook of clinical research ethics. New York: Oxford University Press; 2008. p. 245–60.
18. Miller FG. Recruiting research participants. In: Emanuel EJ, editor. The Oxford textbook of clinical research ethics. New York: Oxford University Press; 2008. p. 397–403.
19. Meltzer LA, Childress JF. What is fair participant selection? In: Emanuel EJ, editor. The Oxford textbook of clinical research ethics. New York: Oxford University Press; 2008. p. 377–85.
20. Agrawal M, Emanuel EJ. Phase I oncology research. In: Emanuel EJ, editor. The Oxford textbook of clinical research ethics. New York: Oxford University Press; 2008. p. 356–66.
21. Bowen AJ. Models of institutional review board function. In: Emanuel EJ, editor. The Oxford textbook of clinical research ethics. New York: Oxford University Press; 2008. p. 552–9.
22. Kant I. Groundwork for the metaphysics of morals (1785). New Haven: Yale University Press; 2002.
23. Capron AM. Legal and regulatory standards of informed consent in research. In: Emanuel EJ, editor. The Oxford textbook of clinical research ethics. New York: Oxford University Press; 2008. p. 613–32.
24. Wendler D. The assent requirement in pediatric research. In: Emanuel EJ, editor. The Oxford textbook of clinical research ethics. New York: Oxford University Press; 2008. p. 661–9.
25. Karlawish JHT. Emergency research. In: Emanuel EJ, editor. The Oxford textbook of clinical research ethics. New York: Oxford University Press; 2008. p. 280–9.
26. Lemmens T. Conflict of interest in medical research. In: Emanuel EJ, editor. The Oxford textbook of clinical research ethics. New York: Oxford University Press; 2008. p. 747–57.

27. Gelsinger PL. Uninformed consent: the case of Jesse gelsinger. In: Lemmens T, Waring DR, editors. Law and ethics in biomedical research: regulation, conflict of interest, and liability. Toronto: University of Toronto Press; 2006. p. 12–32.
28. Steinbrook R. The gelsinger case. In: Emanuel EJ, editor. The Oxford textbook of clinical research ethics. New York: Oxford University Press; 2008. p. 110–20.
29. Albert T, Wager E. How to handle authorship disputes: a guide for new researchers. In: White C, editor. The COPE report 2003: annual report of the Committee on Publication Ethics. London: Wiley; 2004. p. 32–4. Available from: http://publicationethics.org/files/u2/2003pdf12.pdf. Cited 18 Dec 2012.
30. Kim SY, Millard RW, Nesbit P, Cox C, Caine ED. Potential research participants' views regarding researcher and institutional financial conflicts of interest. J Med Ethics. 2004;30:73–9.
31. Angelos P, Murphy TF, Sampson H, Hollings DD, Kshettry V. Informed consent, capitation, and conflict of interest in clinical trials: views from the field. Surgery. 2006;140:740–8.
32. International Committee of Medical Journal Editors. Uniform Requirements for Manuscripts Submitted to Biomedical Journals: Ethical Considerations in the Conduct and Reporting of Research: Authorship and Contributorship. 2005. Available from: http://www.icmje.org/ethical_1author.html. Cited 18 Dec 2012.
33. Office of Science and Technology Policy, Executive Office of the President. Federal Policy on Research Misconduct; Preamble for Research Misconduct Policy; Notification of Final Policy. Fed Regist. 2000;65(235):76260–4.
34. Resnik DB. Fraud, fabrication, and falsification. In: Emanuel EJ, editor. The Oxford textbook of clinical research ethics. New York: Oxford University Press; 2008. p. 787–94.

Chapter 14
Trial Design: Overview of Study Designs (Phase I, II, III, IV, Factorial Design)

Brigid K. Killelea and Anees B. Chagpar

14.1 Introduction

Fundamental to all intellectual inquiry is a scientific method for answering research questions. For the majority of medical and surgical questions that seek to determine if one intervention is better than another, clinical trials are at the cornerstone of all research studies. These, however, must be well designed, with well-thought-out methods and controls such that relevant answers can be derived to fundamental questions.

14.2 The Basics

All research starts with a fundamental question, for which the investigator (and likely the rest of the medical/scientific community) is at equipoise. Is treatment A better than treatment B? It is critical at the outset to define the research questions: What is the intervention to be tested? What is the alternative? What are the outcome measures? What are potential confounders that need to be taken into consideration?

After defining the question, one needs to address feasibility. What is the population to be studied? How rare or common is the condition? How many patients will be needed to answer the research question in a robust fashion? It is important to engage a statistician early in the process as they can help to better design the study but can also determine a sample size based on the research question that can then determine the feasibility of the study planned.

Inclusion and exclusion criteria are critical to define. On the one hand, having broad inclusion criteria may allow for more external generalizability; however, it may also limit the applicability of the trial.

B.K. Killelea • A.B. Chagpar (✉)
Department of Surgery, Yale University School of Medicine, New Haven, CT, USA
e-mail: anees.chagpar@yale.edu

T.M. Pawlik, J.A. Sosa (eds.), *Success in Academic Surgery: Clinical Trials*,
Success in Academic Surgery, DOI 10.1007/978-1-4471-4679-7_14,

Next, one needs to take into consideration the regulatory issues regarding clinical trials. All human subjects research must be conducted in an ethical manner, employing good clinical practices. Regulatory guidelines for the conduct of clinical trials in medicine have since been established and are documented in the Nuremburg Code of 1947 [1] and the Declaration of Helsinki of 1964 [2]. These regulations were designed to protect subjects and require informed consent from all study participants. Therefore, at the outset, one needs to consider how one will protect personal health information and privacy of patients, how data will be stored, etc. For clinical trials, one needs to be cognizant of risks and benefits and how these are explained to patients in an informed consent process. Respect for autonomy and the ability for patients to withdraw from studies also needs to be taken into account when considering sample size calculations. Finally, it is critical that all studies be reviewed and approved by an appropriate institutional review board.

14.3 Phases of Trials

Traditionally, clinical trials go through several phases, which are required particularly for drugs to obtain approval through the Food and Drug Administration or other regulatory bodies. These are enumerated below and shown in Fig. 14.1.

14.3.1 Phase 0

Phase 0 trials are preclinical trials and are intended to expedite the clinical evaluation of new molecular entities. They are often performed in animals. The aim is to look for a dose–response pattern and gather information to guide transition into phase I.

14.3.2 Phase I

After demonstrating safety in animals, the next step is to document safety in humans. Stage I trials are usually small, and the study population may be comprised of volunteers who are in good health or those for whom there is no effective treatment (e.g., metastatic disease). Phase I trials aim to establish safety and perform pharmacokinetic studies that help to determine maximum tolerated dose.

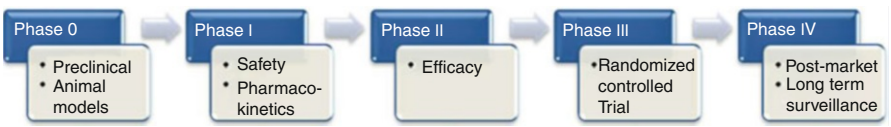


Fig. 14.1 Phases of clinical trials – 0, 1, 2, 3, 4

14.3.3 Phase II

The purpose of phase II trials is to determine efficacy, i.e., whether the treatment at a particular dose/regimen is effective in ideal situations. They determine therapeutic activity and further evaluate toxicity and/or side effects.

14.3.4 Phase III

After demonstrating that a given drug is safe and has a reasonable chance of improving or affecting a disease, phase III trials are undertaken to demonstrate a new drug or treatment is superior to current standard of care. These are usually randomized controlled clinical trials, with large numbers of patients. If only one drug is being evaluated, the design is often placebo controlled and done in parallel. If two or more drugs are being evaluated, a factorial or crossover design is also possible [3]. See Fig. 14.2.

- *Parallel-group design* is a simple design in which each participant is randomly assigned to receive a particular treatment or not. Multiple drugs or treatments may be tested versus each other, and there is often a control or placebo group. Different doses of the same drug can also be simultaneously evaluated. This design is simple and eliminates the potential for drug interaction.
- In a *crossover design*, in general, two treatments are evaluated. Each group receives two treatments but the order is randomly assigned. Prior to the crossover, the design appears to be similar to a parallel-group design. However, the crossover allows one to determine the benefit of adding one drug/treatment to another in a sequential fashion and which order is most beneficial. This design is appropriate for chronic conditions that are stable over time and for interventions that last a short time. It is important that the treatment drugs do not react with each other to avoid making inaccurate conclusions about the results. It should be noted that crossover designs have also been used to evaluate timing of surgical treatments; for example, in the NSABP B-18 trial, patients were randomized to receive either (neoadjuvant) chemotherapy followed by surgery or surgery followed by chemotherapy in order whether one sequence was superior to the other.
- In a *factorial design*, two drugs or interventions can be simultaneously evaluated. With two drugs, four combinations of treatments and placebo are possible. Patients are randomly assigned to each group. For example, one group might receive drug A and drug B. Another group would receive drug A and placebo. Another would receive drug B and placebo, and another would receive two placebos. The factorial design can be very efficient, as data is gathered on two drugs at the same time. A drawback of the factorial design is the concern over potential drug interactions; however, this design also allows for the determination of synergies between two treatments.

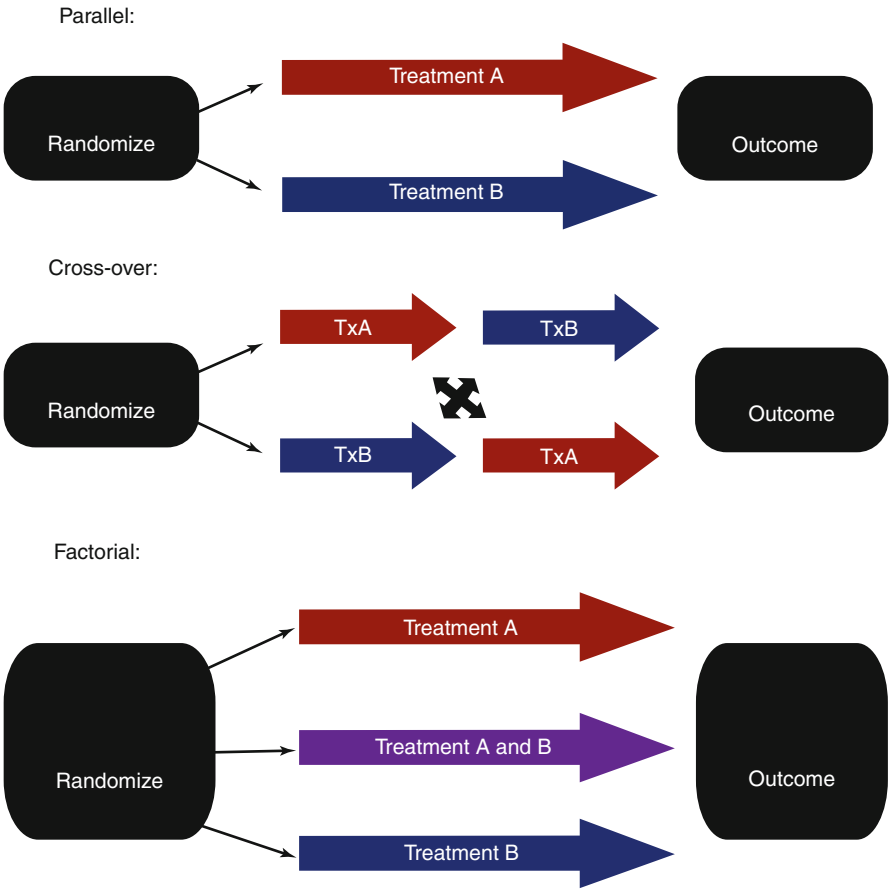


Fig. 14.2 Clinical trial design – parallel, crossover, and factorial

14.3.5 Phase IV

Phase IV studies are post-marketing studies that are conducted after the drug has been approved for use. The purpose of phase IV studies is to gather information regarding additional potential side effects in diverse populations and gather long-term follow-up data.

For surgical trials of new equipment or implants, a parallel series of trials has been proposed [4].

14.3.6 Phase I: Laboratory Study

This phase, similar to phase 0 trials for drug, is preclinical and often involves animal models.

14.3.7 Phase II: Cohort Study

As devices and implants do not have a "dose" that needs to be titrated, this phase is akin to phase II trials where a tightly controlled study is done on a cohort of patients to determine efficacy of a particular device or treatment.

14.3.8 Phase III: Randomized Controlled Trial

As with other phase III randomized controlled trials, the purpose of these studies is to compare outcomes with a particular device, implant, or treatment vs. standard of care.

14.3.9 Phase IV: Surveillance Study

Again, careful evaluation must continue for long-term sequelae after a phase III randomized controlled trial; for devices or surgical implants, this is similar to phase IV trials for drugs.

14.4 Randomization

The gold standard to determine if one intervention is better than another is the randomized controlled trial, which provides the strongest evidence for a cause-effect association between an intervention and an outcome. The process of randomization effectively limits bias, such that differences in outcome may be attributable to differences in treatment groups, rather than to inherent differences or confounding factors between the two groups [5, 6].

Several randomization strategies exist. *Simple randomization* is analogous to a coin toss; heads means the patient will be randomized to the treatment arm, and tails means they will be in the control group. Randomization tables can simply allocate patients to one arm or another in whatever ratio the trial dictates. Another (slightly more complicated) randomization strategy is *block randomization*, where patients are divided into blocks (which are often based on some characteristic) and randomized within each block. This is also called *stratified randomization*, as patients are initially stratified according to a key variable or set of variables (e.g., age, race, sex, stage) and then randomized within each stratum. In surgical trials, stratifying by surgeon has been used to adjust for the experience of individual surgeons [7]. In very large trials, stratification is usually unnecessary, as large imbalances tend to diminish with greater numbers of subjects.

While often patients are randomized equally between groups, this need not be the case. Unequal randomization where patients are randomized 2:1, 3:1, etc., are associated with less power but may have other advantages.

In addition, while patients are often randomized at the individual level, larger-scale trials may utilize a *cluster design*, in which predefined groups (e.g., members of a particular clinic, region) receive one treatment and are compared to a similar group who receives another. This design may be chosen if the treatment or intervention can only be administered on a large scale, such as a behavior modification intervention or some screening initiatives [8]. Sophisticated statistical techniques need to account for differences between clusters and the fact that subjects within a cluster are correlated.

14.5 Blinding

Ideally, both patients and study personnel are blinded, meaning that they are unaware which patients are receiving active drug or placebo. This is referred to as a *double-blind* study. In a *single-blind* study, either the investigator or the participant knows which group is receiving active treatment, and in an *unblinded* trial, both investigators and patients are aware of treatment. There are times when it is not feasible to blind patients or investigators to treatment arm, particularly in some surgical studies.

14.6 Hypothesis Testing and Data Analysis

A different way to categorize randomized controlled clinical trials is by hypothesis, and the criteria by which the null hypothesis will be rejected. In a *superiority* trial, the hypothesis is that one drug or treatment is better than the other. In a *non-inferiority* trial, the hypothesis is that a new drug or treatment is no worse than the existing drug. *Equivalence* trials are based on the hypothesis that there is no difference between two drugs or treatments. This distinction is important in constructing the trial, in calculating sample size needed, and in interpretation of the results.

It is also important to define the goals and objectives of the study and the outcome variables before commencing the trial. There should be a clear primary aim or question the trial seeks to answer. Any secondary aims or hypotheses need also to be identified. *Subgroup analyses*, which refers to the practice of dividing the study participants into specific groups and doing additional analyses, need to be specified at the outset. Subgroup analyses are observational by definition and are not based on randomized groups.

In general, an overall analytic plan should be prespecified prior to starting a trial. Often, an *intention-to-treat* analysis is utilized, where patients who were randomized to each group are compared, regardless of whether or not they completed the specified treatment. In this way, “effectiveness” or how a treatment would work in real-life circumstances can be evaluated. Alternatively, analyses may be done in *per-protocol* manner, where only patients who completed the specified treatment

are compared. In this way, "efficacy" or how a treatment will work in ideal circumstances can be assessed. One should discuss with a statistician how missing data will be managed, as there are a number of specific statistical methods to manage incomplete data fields, and the fact that a proportion of patients may either drop out of the study or be lost to follow-up should be taken into consideration when computing sample size.

14.7 Interim Analyses and Bayesian Design

When planning a clinical trial, it is often necessary to plan interim analyses to assess both toxicities and adverse effects, as well as whether one arm is clearly superior to the other – both of which may require a trial to stop prematurely. Significant adverse events must be reported promptly to institutional review boards, but an interim analysis will catalog all adverse events such that patterns may become clear. In addition, "stopping rules" that would mandate a trial close prematurely when one arm is found to be significantly better than the other (and therefore it would be unethical to continue to accrue to the poorer arm) should be set in advance. An independent data safety and monitoring board should be identified prior to the study commencing. This body would meet periodically to review the study's progress and discuss the findings of interim analyses, keeping the primary investigator and key study personnel blinded until the study conclusion.

One of the key disadvantages of randomized trials is that often they require large numbers of patients and significant follow-up before a conclusion is reached about the superiority of one treatment over another. At the same time, newer treatments are being developed, and incorporation of these agents may be more difficult in trials that are "fixed" by the number of patients to be accrued on a given treatment arm. Some modern trials have now incorporated a "Bayesian approach" whereby as we learn about the effectiveness of certain treatments in given subgroups of patients, they are compared to the next treatment regimen [9, 10]. This approach is analogous to doing a series of interim analyses and changing the arms of the parallel-group design according to the results. Examples include the i-SPY and BATTLE trials, where different treatments are compared given various molecular subtypes [11, 12].

14.8 Correlative Science and Other Analyses

Clinical trials are often expensive endeavors that consume a significant amount of both clinical and human resources. Much of this is worth the investment, as the data obtained can answer important questions in a well-defined population for whom treatments are standardized. In order to maximize the knowledge that can be gained from clinical trials, thought should be put into other potential questions that can be answered in the context of the trial. Correlative science, quality of life surveys, and

other analyses are commonly built into trials. It is important to consider potential secondary studies that could be done in the context of a trial up front, such that data needed to answer these questions are collected prospectively as part of the trial.

14.9 Reporting Trial Design

The Consolidated Standards of Reporting Trials (CONSORT) Statement is a guide for investigators in reporting randomized controlled trial in such a way as to optimize the readers' ability "to understand a trial's design, conduct, analysis, interpretation, and to assess the validity of its results." [13] CONSORT has a checklist that can ensure that trial designs and methods are appropriately reported. A similar 10-item Checklist to Evaluate A Report of Non-Pharmacologic Trials (CLEAR NPT) was specifically developed for non-pharmacologic treatments, which may be more relevant for some surgical studies [14]. As surgeons, it is critical that meticulous attention is paid to the details of designing trials so that these may withstand the rigor of critical assessment [15].

14.10 Summary

Clinical trials remain the "gold standard" for most clinical research, and while surgical trials have intricacies that are unique, adherence to fundamentals of proper design and conduct of clinical trials remain mandatory for key insights to be obtained.

References

1. The Nuremberg Code. JAMA. 1996;276:1691.
2. World Medical Association declaration of Helsinki. Recommendations guiding physicians in biomedical research involving human subjects. JAMA. 1997;277:925–6.
3. Gurusamy KS, Gluud C, Nikolova D, Davidson BR. Design of surgical randomized controlled trials involving multiple interventions. J Surg Res. 2011;165:118–27.
4. Gross M. Innovations in surgery. A proposal for phased clinical trials. J Bone Joint Surg Br. 1993;75:351–4.
5. Sabatine MS. Cardiology patient page. Randomized, controlled trials. Circulation. 2011;124:e832–4.
6. Stanley K. Design of randomized controlled trials. Circulation. 2007;115:1164–9.
7. Fung EK, Lore Jr JM. Randomized controlled trials for evaluating surgical questions. Arch Otolaryngol Head Neck Surg. 2002;128:631–4.
8. Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. Contemp Clin Trials. 2007;28:182–91.

9. Lee JJ, Xuemin G, Suyu L. Bayesian adaptive randomization designs for targeted agent development. Clin Trials. 2010;7:584–96.
10. Lee JJ, Chu CT. Bayesian clinical trials in action. Stat Med. 2012;31:2955–72.
11. Barker AD, Sigman CC, Kelloff GJ, Hylton NM, Berry DA, Esserman LJ. I-SPY 2: an adaptive breast cancer trial design in the setting of neoadjuvant chemotherapy. Clin Pharmacol Ther. 2009;86:97–100.
12. Kim ES, Herbst RS, Wistuba II, et al. The BATTLE trial: personalizing therapy for lung cancer. Cancer Discov. 2011;1:44–53.
13. Schulz KF, Altman DG, Moher D, for the CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. Ann Int Med 2010;152. Epub 24 March.
14. Boutron I, Moher D, Tugwell P, et al. A checklist to evaluate a report of a nonpharmacological trial (CLEAR NPT) was developed using consensus. J Clin Epidemiol. 2005;58:1233–40.
15. Jacquier I, Boutron I, Moher D, Roy C, Ravaud P. The reporting of randomized clinical trials using a surgical intervention is in need of immediate improvement: a systematic review. Ann Surg. 2006;244:677–83.

Index

T.M. Pawlik, J.A. Sosa (eds.), *Success in Academic Surgery: Clinical Trials,* Success in Academic Surgery, DOI 10.1007/978-1-4471-4679-7,