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

Timothy M. Pawlik
Julie A. Sosa *Editors*

Clinical Trials

Second Edition

Springer

Success in Academic Surgery

Series editors:

Lillian Kao
McGovern School
The University of Texas Health Science Centre
Houston, TX
USA

Herbert Chen
Department of Surgery
University of Alabama at Birmingham
Birmingham, AL
USA

More information about this series at http://www.springer.com/series/11216

Timothy M. Pawlik • Julie A. Sosa
Editors

Clinical Trials

Second Edition

Springer

Editors
Timothy M. Pawlik
Chair, Department of Surgery
The Urban Meyer III and Shelley Meyer Chair for Cancer Research
The Ohio State University Wexner Medical
Columbus, OH
USA

Julie A. Sosa
Leon Goldman, MD Distinguished Professor of Surgery and Chair
Department of Surgery
University of California, San Francisco
San Francisco, CA
USA

Professor, Department of Medicine
University of California, San Francisco
San Francisco, CA
USA

ISSN 2194-7481 ISSN 2194-749X (electronic)
Success in Academic Surgery
ISBN 978-3-030-35487-9 ISBN 978-3-030-35488-6 (eBook)
https://doi.org/10.1007/978-3-030-35488-6

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

Chapter 1
The History of Clinical Trials

Janice Hu, Justin Barr, and Georgia M. Beasley

1.1 Introduction

A clinical trial is a purposeful comparison of medical interventions, including placebos, against one another to determine the safest, most efficacious means of treating pathology. The history of clinical research in surgery sheds light on both the successes and the challenges that academic surgeons faced when developing therapies for their patients.

1.2 Early History

Clinical trials had little role in the ancient world where accepted disease theories rendered them all but irrelevant. In many older cultures, disease and healing were perceived to stem from supernatural and divine forces. In Greece during the fifth century BCE, patients sought healing through "incubation," or sleep, in temples of the healing god Asclepius. It was around this time that a new form of medicine arose, marking a major innovation in the treatment of disease. Unlike supernatural theories, Hippocrates' method involved seeking the cause of illness in natural

J. Hu
Duke University School of Medicine, Durham, NC, USA

J. Barr
Duke Department of Surgery, Durham, NC, USA

G. M. Beasley (✉)
Duke University School of Medicine, Durham, NC, USA

Duke Department of Surgery, Durham, NC, USA
e-mail: georgia.beasley@duke.edu

T. M. Pawlik, J. A. Sosa (eds.), *Clinical Trials*, Success in Academic Surgery,
https://doi.org/10.1007/978-3-030-35488-6_1

factors involving the composition of the body's humors. An oeuvre of texts known as the *Hippocratic Corpus*, written by numerous authors over many decades until the first half of the fourth century BCE, established that physicians could learn through observations and actions. Yet the ancient Greeks did not perform clinical trials to test their hypotheses. Moreover, the highly individualized understanding of disease made broadly applicable treatments rare, vitiating the value of clinical trials. The Greek had "freed himself of religion to become the prisoner of philosophy" [1]. This dogma largely continued through the Roman world [2].

In 1025 CE, the Persian physician Avicenna wrote the widely used medical treatise *The Canon of Medicine* in which he laid down a precise guide for empirical investigation of the effectiveness of medical drugs and substances [3]. He recommended studying two cases of contrary types, along with the timing and reproducibility of drug effects so that consequence and accident are not confused. Moreover, he advocated for experimentation on the human body, since testing a drug on a lion or a horse might not prove anything about its effect on man. The pharmacology discussed in Avicenna's treatise was used extensively in medical schools across Europe as late as 1650 [4]. Although Avicenna advocated for the empirical study of drugs, his *Canon* did not lead to the widespread engagement of experiments and empiricism. Instead, the Medieval Era (800–1400 CE) was characterized by textual dependence and interpretation that prized the authority of the ancients over experimental evidence [5]. Moreover, while extant sources such as the *Hippocratic Corpus* and the *Canon* defined elite, academic-based medicine, the vast majority of medical care was delivered by untrained, unlicensed, and irregular practitioners, most of whom were illiterate. This practice went largely unrecorded and likely relied on a combination of superstition, tradition, and empiricism.

1.2.1 Early Modern Era (1500–1800)

With the dawn of the early modern era in the sixteenth century, there was a general intellectual shift away from dogmatic textual dependence and toward empirical investigation. This was evident in multiple arenas including heliocentric theories of astronomy put forth by Nicolaus Copernicus, anatomical observations made by Andreas Vesalius, and navigational feats like those by Christopher Columbus. They also appeared in medicine.

One of the first clinical trials was accidentally conducted in 1537 by the French surgeon Ambroise Paré when he ran out of the boiling oil that was conventionally used to treat bullet wounds and resorted to giving some soldiers a balm made from egg yolks, rose oil, and turpentine [6]. He awoke the following morning to find that patients who received the new treatment were resting well with little discomfort and swelling, whereas those who were cauterized with oil were "feverish with much

Fig. 1.1 Ambroise Paré et l'examen d'un malade [Ambroise Paré examining a patient] by James Bertrand (1823–1887), from the Charles de Bruyères Museum collection in Remiremont. (*Source*: Ji-Elle, license CC:BY-SA) [8]

pain and swelling about their wounds." Reflecting on this experience, he noted "I resolved with myself never more to burn thus cruelly poor men wounded with gunshot" [7]. This observation, widely published, changed clinical practice as military surgeons across Europe began to eschew boiling oil in favor of less painful remedies (Fig. 1.1).

Systematic tests of disease management tackled the pre-fifteenth century Galenic tradition of wound management, characterized by gradual "wet healing" that involved forcing wounds open and applying emollients. This conventional method often led to poor outcomes. From 1580 to 1583, Spanish surgeon Bartolomé Hidalgo de Agüero challenged this notion by examining hospital records, finding that his own method of "dry healing"—cleaning the wound with white wine, removing damaged tissue, bringing the edges together, applying drying compounds, and covering the wound with a bandage—led to a far lower mortality rate compared to the Galenic technique [9].

The trend of empiricism continued to grow as physicians set forth hypotheses and began testing them through observation. Paré and Agüero belonged to a group of sixteenth century practitioners who were willing to trust their observations and personal experience over ancient traditions and dogma. Yet two centuries would pass before the launch of the first rigorous prospective trial.

Scottish surgeon James Lind randomized six pairs of sailors to different treatments for scurvy in 1747, finding that citrus fruits were the most effective therapy [10]. Despite the soundness of his methods and the irrefragability of his results, his conclusion had little impact on medical opinion in Britain, exposing an ongoing theme through this history: the challenge of even the best clinical trial actually changing medical practice. It ultimately requires many more decades, with thousands of additional deaths, for professional opinion to adopt lemons as a scurvy prophylactic.

1.3 The Emerging Importance of Statistics

Comparative retrospective analyses played an important role in building toward controlled trials in medicine and surgery. Statistics, or the practice of collecting and analyzing large amount of numerical data, emerged as an important tool in treatment evaluation. By the eighteenth century, several case series propelled arguments about the utility, methods, and timing of limb amputations [11, 12]. Lithotomists published numerical evidence on bladder stone removal, debating the merits of lithotripsy compared to lithotomy and examining mortality among age subgroups [13–15]. In the 1820s, Pierre-Charles-Alexandre Louis used his "numerical method" on aggregated clinical data to cast doubt on the practice of bloodletting [16, 17]. Furthermore, statistics featured prominently in debates surrounding perioperative innovations such as anesthesia and Lister's "antiseptic method" of carbolic acid for surgical wounds beginning in 1867 [18, 19]. This portended the clear role and need for stronger evidence to evaluate theories of disease management. It also demonstrated the shift from highly individualized disease states as understood in ancient and medieval medicine to a more ontological notion of sickness where a single intervention had the potential to apply to all patients suffering from the same pathology. This critical theoretical transition made clinical trials relevant. Moreover, as anesthesia and antisepsis allowed surgeons to delve further into internal organs and conduct more elective procedures, there arose a clear need to provide proof of safety and benefit.

1.4 Prospective Clinical Trials Begin

In the nineteenth century, surgeons joined in performing prospective trials by first using nonrandom methods of treatment assignment such as alternate allocation. In perhaps the earliest example of this, an 1816 medical dissertation describes how military surgeons performed a controlled trial on 366 soldiers in the Peninsular War to assess the effects of bloodletting for fever. Although there are uncertainties surrounding the authenticity of this report [20], it nonetheless illustrates the emerging desire among surgeons to control for factors other than the treatment of interest:

> It had been so arranged, that this number was admitted, alternately, in such a manner that each of us had one third of the whole. The sick were indiscriminately received, and were attended as nearly as possible with the same care and accommodated with the same comforts. One third of the whole were soldiers of the 61st Regiment, the remainder of my own (the 42nd) Regiment. Neither Mr. Anderson nor I ever once employed the lancet. He lost two, I four cases; whilst out of the other third *[treated with bloodletting by the third surgeon]* thirty five patients died [21].

The last decades of the nineteenth century witnessed the publication of other prospective surgical studies using alternate allocation. These included catheterization for urethrotomies, capsulotomy following removal of cataracts, and pediatric hernia management [22–24]. The goals of these researchers were twofold: (1) to

make firmer distinctions among different interventions and (2) to demonstrate impartiality. In 1893, W. T. Bull explained how alternate allocation reduced bias when comparing a spring truss to a skein wool truss for the treatment of pediatric hernias:

> In children under the age of 1 year the worsted or so-called 'hank truss' has been extensively tried. This truss has been very highly praised by some, and as strongly condemned by others. During the past year an attempt has been made to give it an impartial trial, and alternate cases up to the age of 1 year were treated by the 'hank' and the light spring truss. The results in 240 cases carefully followed up led us to discard the hank truss as a routine method of treatment, although there are still a few cases—for example, very young and ill-nourished infants where it fills a useful but temporary place [24].

Although prospective studies comparing groups of patients emerged into the professional surgery landscape, individual patient outcomes remained powerful guides for surgical management. After all, Bull still advocates for the use of the hank truss in "a few cases" based on select patient characteristics. In fact, despite the emergence of prospective controlled studies, case series feature prominently in the body of published surgical evidence well into the late nineteenth and early twentieth centuries [25]. Indeed, these retrospective studies helped surgeons select operative techniques amidst the variability of their patient populations. They also led, however, to protracted debates about competing techniques and to the propagation of now-defunct operations including treatments for ptosis, constipation, and autonomic nerve dysfunction [26]. Surgeons tended to publish case series that promoted their own opinions, leading to unresolved debates in areas such as radical mastectomy and prostate surgery [27, 28]. Biased results continued to highlight the need for more carefully designed investigations.

1.5 The First Randomized Clinical Trials

In order for the randomized trial to become the gold standard in guiding medical practice, its various constituents, including controls, blinding, quantification, and randomization, needed to undergo their own evolution [25]. The randomization component in particular is important because it eliminates selection bias, balances treatment groups with respect to confounders, and forms the basis of statistical tests which assume equality of treatments. After R.A. Fisher demonstrated the utility of randomization and novel statistical analysis techniques in agricultural research in the 1920s, researchers began adapting this method in medicine.

The impetus for randomized trials also depended on the interaction between professional interests and the regulatory environment. Scandals surrounding drug safety and the for-profit pharmaceutical industry in the 1930s prompted clinical trials in medicine. Journalists, consumer protection organizations, and federal regulators began mounting a campaign for stronger regulatory authority by publicizing a list of harmful products including radioactive beverages and ineffective "cures" for diabetes and tuberculosis [29]. The Food and Drug Administration (FDA) began

to require random assignment and control groups in pharmaceutical testing. As surgeons were relatively unaffected by these controversies, they enjoyed greater freedom in the early twentieth century to "adopt, adapt, or invent" through personal experience and case studies [28]. After all, surgery physically rearranges body tissues, and the end product is visible proof that an intervention has taken place. A purported "magic pill" is much more vulnerable to skepticism.

In 1931, American researchers published an article in the *American Review of Tuberculosis* depicting the first randomized controlled trial with blinding and placebo controls. Amberson and his colleagues used a coin flip randomizing tuberculosis patients to receive either sanocrysin (a gold compound) or distilled water. The resulting data demonstrated that all of the patients receiving sanocrysin suffered adverse systemic drug effects, with no evidence of therapeutic benefit at follow-up [30]. In the very same journal issue, Brock published a study arriving at very different conclusions, that sanocrysin had "an outstanding clinical effect on exudative tuberculosis in white patients," although "very little effect in limiting the progression of disease in black patients" [31]. In comparison to Amberson's trial, Brock's study was demonstrably weaker; he observed 46 patients who were given varying dosages of sanocrysin, did not have an untreated control group, and did not control for baseline differences in treatment setting and disease stage between black and white patients [32]. Clinicians recognized that Amberson's randomized controlled trial provided stronger evidence, and thus gold therapy for tuberculosis fell into disrepute throughout America.

The first multicenter trials addressing the treatment of pulmonary tuberculosis with streptomycin were published in the United Kingdom in 1948 and the United States in 1952. The British study included 107 patients from 7 centers and concluded that streptomycin-treated patients experienced significantly better outcomes compared to control patients. The Veterans Administration and the United States Armed Services added to the body of evidence from multicenter trials, with good success [33].

One of the earliest randomized controlled trials related to surgery was anesthesiologist Henry Beecher's 1955 investigation of three different anti-emetics for postoperative vomiting [34]. The year 1958 saw the launch of several randomized controlled trials on surgical procedures, including the management of upper gastrointestinal bleeding, prophylactic surgery for esophageal varices, internal mammary artery ligation, and radical mastectomy [35–38]. Perhaps the largest and most well-known early randomized controlled trial in surgery was performed by C. Goligher's team in Leeds and York in 1959. The study randomized 634 carefully selected patients to one of three operations for duodenal ulcers. This trial helped lay the design foundation for future trials in surgery. On the importance of random assignment, the study remarks:

> This method of randomization may strike some as very impersonal, but we would point out that during the time the trial has been in progress surgical opinion throughout the country on the choice of elective operation for duodenal ulceration has been so divided that in any large hospital several different methods were already in use. Which one would be performed on an individual patient has depended largely on the personal predilection of the particular

> surgeon to whom he happened to be referred and not on any accurate knowledge of the relative late results. Our trial has merely organized somewhat this pre-existing system of random usage in order to extract more reliable information from it [39].

Randomized trials aimed, therefore, to settle the conflict of "divided" opinions in the country regarding surgical management of specific diseases and to set a precedent of basing treatments on proven effectiveness rather than on individual surgeon preference. Yet the opening line of the quote clearly articulates how foreign and potentially controversial this methodology was to surgeons in the 1950s and 1960s, many of whom questioned the ethics of denying patients the treatment perceived to be most efficacious.

Despite the lack of "high-quality evidence" provided by carefully designed investigations, procedures such as vagotomy and subtotal gastrectomy, among many others, came to be regularly practiced. How did this occur? And how did the standard of proof transform from expert opinion to more standardized trials? A focused history of clinical research in breast cancer surgery offers a lens through which to understand this phenomenon.

1.6 The History of Clinical Research in Breast Cancer Surgery

Propelled by the theory that breast cancer spread centrifugally in the plane of subcutaneous tissues and lymphatics, radical mastectomy remained a mainstay of surgical treatment throughout the first half of the twentieth century [40]. William Halsted did much to pioneer the radical mastectomy, performing the first "Halsted mastectomy" in 1882 [41]. He used clinical and pathologic findings from a series of 210 cases, of which he marked 42% as 3-year cures, results that surpassed those of other surgeons at the time [42]. A 1924 review of 20,000 cases of breast cancer by British statistician Janet Lane-Claypon reported that radical mastectomy offered 43.2% three-year survival rates compared to less than 30% survival from more conservative operations [43]. Halsted's operation peaked in popularity after World War II as the American Cancer Society pushed for early detection and removal of breast cancers. Some surgeons, believing Halsted's operation to be insufficient, pushed the envelope even further through "superradical" operations such as removing ribs, deep lymph nodes, limbs, and even internal organs to eradicate cancer cells [27] (Fig. 1.2).

Several case series in Europe began to cast doubt on this prevailing theory, reporting that, in stage I and II breast cancers, more conservative operations resulted in similar survival rates compared to radical mastectomy [45–47]. Similarly, a few American physicians such as Barney Crile presented retrospective data indicating that less radical procedures resulted in equal or better results with fewer side effects compared to the Halsted approach [48]. Moreover, radiotherapy pioneered at the Curie Institute in Paris emerged as a new modality for treatment, and case series demonstrated that when it was either used alone or in combination

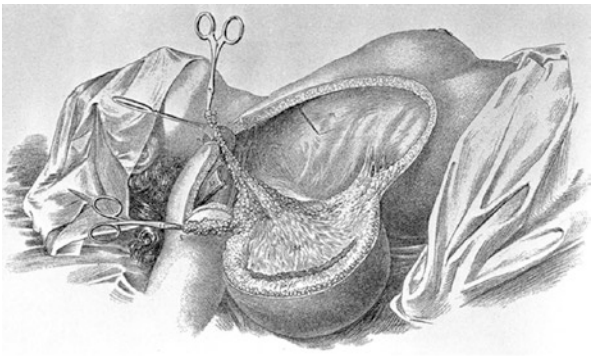


Fig. 1.2 Original drawing of the radical mastectomy reported by William S. Halsted in 1894. (*Source*: William Stewart Halsted, Surgical papers, Wellcome Collection) [44]

with more conservative surgeries, radiation appeared to lend similar or better results compared to radical mastectomy [49, 50]. Physician-historian Barron Lerner points out the strength of a "natural experiment" by Smith and Meyer which showed that simple mastectomies performed during World War II due to staff shortage—therefore representing patients with similar disease severity compared to patients treated in peacetime—led to similar results between simple and radical mastectomies [27, 51].

Radical mastectomy remained standard of care, however. It was viewed as unethical to deprive patients of the ostensibly superior Halsted radical procedure. Surgery carried a strong culture of reliance on expertise gained from firsthand operative experience rather than on biostatistics. Even proponents of conservative surgery such as Barney Crile did not advocate for randomized trials; they felt their personal operative records held sufficient proof of the merits of their approach. Surgeons were concerned that randomized trials would impede on their authority to make individualized decisions for their patients [27].

As providers in the United States debated whether to perform randomized controlled trials to study breast cancer surgery, these very trials were initiated in Europe. Beginning in 1951, researchers used alternate allocation to compare simple mastectomy and radiation with radical mastectomy, showing that the more radical procedure afforded no additional survival benefit [52]. In 1958, radiotherapists Diana Brinkley and J. L. Haybittle launched a randomized controlled study in Cambridge comparing simple mastectomy to radical mastectomy, with all patients receiving radiotherapy. Five- and ten-year survival was equivalent between the two groups [36].

Despite the apparent need for more rigorous studies, it was not until 1971 that the first randomized controlled trial on breast cancer surgery began in the United States. Bernard Fisher began enrolling breast cancer patients to compare radical mastectomy with simple mastectomy [53]. At 25-year follow-up, the study found there was no significant survival advantage gained from performing radical mastectomy to remove occult positive nodes at the time of initial surgery or from radiation therapy. These findings further supported the notion that outcomes from breast cancer surgery relied not on radicality but rather on adequate control of local disease and treatment of secondary tumor spread.

Physician-historian David S. Jones points out that RCTs often were not required or even relevant to promulgate changes in surgical practice [25, 38]. The rates of radical mastectomy had already fallen from 50% in 1972 to 3% in 1981, long before the publication of Fisher's trial [54]. Operative management of breast cancer was already shifting toward more conservative methods; therefore, the RCT made its impact in tandem with other factors such as patient empowerment and new understandings of disease models [27, 54]. The history of breast cancer surgery research illustrates the evolution from empiric clinical gestalt informing decisions to the use of rigorous trials to support or refute longstanding theories. Randomized controlled trials were not foundational to the move away from radical mastectomies, however, and historically such trials have not shaped surgical practice nearly to the same extent as they have shaped medicine.

1.7 Challenges in the Uptake of RCTs in Surgery

Randomized evaluations of surgical techniques are rare, and many interventions have been widely adopted without rigorous evaluation. In the 1990s, an estimated one half of interventions in internal medicine were based on evidence from RCTs, compared to fewer than 25% of surgical interventions [55–57]. In the latter half of the twentieth century, only 3.4% of all articles in the leading surgical journals were randomized controlled trials [58]. This gradually rose to an estimated 10% by 2006 [59, 60].

There are several reasons for the large-scale delay in the uptake of RCTs in surgery [25, 38]. One major reason stems from the blurred lines between clinical practice, innovation, and research, a phenomenon explored by scholars of surgical history. Sally Wilde and Geoffrey Hirst describe how early twentieth century surgeons constantly combined theories about the body with empirical observations in the operating room to innovate new techniques [61]. Surgeries are not controlled by a regulatory body such as the FDA and can be performed without first undergoing extensive evaluation; therefore, regulatory factors are not an impetus to devote the funding and institutional organization required to support large-scale randomized controlled trials [62]. In a survey of surgeons who had published papers describing innovative surgeries, Reitsma and Moreno found that 14 of 21 surgeons confirmed that their work was research, but only 6 had sought IRB approval, and only 7 mentioned the innovative nature of the procedure in the informed consent document [63]. These findings demonstrated a clear need for education and possibly some minimal criteria that define experimentation in performing surgical procedures. In 2009, the IDEAL Collaboration endorsed several suggestions geared toward improving the assessment of surgical innovations, including the use of prospective databases and registries as well as increasing the number of prospective studies with adequate statistical control techniques [64].

Skeptics also felt that surgery was inherently not amenable to standardization, particularly in comparison to medication, where pills maintain the exact same

chemical composition and dose throughout a trial. In contrast, operations comprise hundreds of steps that individual surgeons continually refine and innovate for each particular patient in the hopes of achieving better outcomes [65]. Unlike a clinical trial testing a new medication, variation in surgical skill and experience will allow some surgeons to achieve an adequate result more quickly, whereas other surgeons may need to perform the procedure multiple times to attain the same results.

The ethics of randomized controlled trials in surgery also carries complexities; for instance, establishing evidence using randomized controlled trials would not be ethical in some procedures due to the risk of harm in the nonoperative group. Moreover, studies with placebo sham surgeries have been viewed as unethical because the benefits cannot outweigh the risks of an invasive procedure [66].

1.8 Ethics and Regulation of Clinical Trials

The development of ethical standards with respect to medical experimentation has been an ongoing concern [67]. Military surgeon Walter Reed utilized some of the first written informed consents (in English and Spanish) for his yellow fever trials in Cuba at the turn of the century [68]. The Nuremberg Code of 1949, issued in reaction to Nazi experimentation, was the first document to set out ethical principles based on informed consent. These principles were revised and released by the World Medical Association in 1964 as the Declaration of Helsinki [69]. When thousands of children were born with birth defects as a result of pregnant women taking the drug thalidomide for morning sickness, the 1962 Kefauver-Harris Amendment to the Food, Drug, and Cosmetic Act set forth legal requirements for "adequate and well-controlled investigations" prior to a drug's approval by the FDA [29, 70]. In 1966, as Henry Beecher was about to publish his exposé on unethical clinical research practices, the US Surgeon General requested that hospitals and universities establish review boards [71].

One of the most infamous clinical trials where ethical principles were lacking was the "Tuskegee Study of Untreated Syphilis" conducted by the US Public Health Service from 1932 to 1972 [72, 73]. It involved nearly 400 black men with late-stage syphilis. When penicillin was found to be an effective cure for syphilis in 1946, the subjects enrolled in the study were not offered this treatment and were not informed of their diagnosis. Jean Heller of the Associated Press broke the story of this study in 1972, revealing that the trial did not have a formal protocol. The magnitude of the risks taken with the subjects involved led many to believe that the Public Health Service had "played" with human lives [72]. The Tuskegee study performed a key role in creating institutions and practices that govern the use of human volunteers in US biomedical research today, but it also introduced a level of distrust between patients and physicians and made the public wary of participating in clinical studies.

In the wake of the tragic disregard for ethical principles in the Tuskegee Study, the National Research Act was signed into law in 1974, culminating in the creation

of the Belmont Report. It put forth three basic ethical principles: respect for persons (to protect autonomy as well as those with diminished autonomy), beneficence (to maximize possible benefits and minimize possible harm), and justice (to divide benefits and burdens of research equally among individuals) [69, 74]. 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.

To date, compared to medical therapies and devices, new surgical techniques have arguably escaped the same type of scrutiny imposed by the FDA and Institutional Review Boards (IRBs) [75]. Designation of a surgical innovation as "experimental" has largely been left to the discretion of the surgeon. To address the concerns that potentially harmful operations could be developed without rigorous evaluation, the American College of Surgeons formulated guidelines in 1995 for the evaluation and application of emerging procedures, urging that new technologies require earlier and continued IRB review, scrutiny of the research protocol, along with a thorough description of informed consent of subjects [63, 76].

The history of clinical research in surgery sheds light on the tension between innovation and strict control. Not long ago, it was common practice for surgeons to take novel operations and technology and apply them to patients after minimal study. The gold standard for building evidence has now assumed the form of a rigorous, expensive, multi-year process. Spurred appropriately by the desire to protect patients from unethical conditions, researchers have foregone rapid innovation in favor of safety. As surgeons navigate ways to improve surgical care, the scientific community will continue to reevaluate the balance between innovation and regulation.

1.9 Recent Times

As clinical trials became more complex, they required additional regulation and administration. Clinical trials at academic centers often ran from specific medical departments. Clinical trial offices (CTOs) emerged over the last two decades to encompass 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 of functions designated to the CTOs across institutions; some were gatekeepers on all budgeting and billing, and others provided educational or liaison services, while still others wielded monitoring and auditing responsibilities for compliance [77]. This review points to the challenge that institutions face when defining the structure of clinical trial administration. CTOs will become increasingly important as there is added pressure on academic organizations to focus their billing and compliance activities, increase communication between researchers, consolidate education and training, decrease costs and infrastructural redundancy, and increase visibility of trials.

To validate their results and share resources, randomized trials need collaboration and organization among multiple institutions. The Early Breast Cancer Trialists' Collaborative Group was a multinational effort to compile results from randomized trials of adjuvant endocrine and cytotoxic treatments [78]. The creation of the National Cancer Institute (NCI) in 1937 designated the start of federally sponsored medical research in the United States and developed into the National Institutes of Health in the post-World War II years. With the aim of facilitating cross-institutional collaboration, the NCI created several cooperative cancer research groups including the National Surgical Adjuvant Breast and Bowel Project. Collaborative research arising from these groups has helped demonstrate that breast conservation surgery is often better than radical mastectomy, for instance [79]. These multicenter trials are attractive for a number of reasons, including large-scale patient recruitment to achieve the needed numbers, attention to regulatory and ethical issues, as well as marketing strategies. Mega-trials often include thousands of patients with significant heterogeneity in demographics, clinical characteristics, comorbidities, and associated therapies. 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 and clinical associations [80]. For these reasons, the enthusiasm for international mega-trials has waned somewhat, and researchers have started to focus their energies in more individualized patient-centered research—not all that dissimilar from the patient-particular practices of Hippocratic physicians 2000 years ago. While there may never be a "perfect" trial, clinicians and researchers will continue to employ old, new, and yet-to-be-invented modalities to test the therapeutic potential of interventions in the everlasting goal of providing the best possible care to patients.

References

1. Majno G. The healing hand: man and wound in the ancient world. Boston: Harvard University Press; 1975.
2. Nutton V. Ancient medicine. New York: Routledge; 2004.
3. Flannery M. Avicenna [updated 12 Oct 2018]. https://www.britannica.com/biography/Avicenna. Accessed 5 Dec 2018.
4. Sirasai N. Avicenna in renaissance Italy: the canon and medical teaching in Italian universities after 1500. Princeton: Princeton University Press; 1987.
5. Siraisi NG. Medieval and early renaissance medicine. Chicago: University of Chicago Press; 1990.
6. Packard F. The life and times of Ambroise Paré. 2nd ed. New York: Paul B. Hoeber; 1925.
7. Paré A. The apologie and treatise of Ambroise Paré. Birmingham, Al: Classics Of Medicine Library; 1984.
8. Ambroise Paré et l'examen d'un malade. 2011. https://commons.wikimedia.org/wiki/File:James_Bertrand-Ambroise_Paré.jpg.
9. Solís C. Bartolomé Hidalgo de Agüero's 16th century, evidence-based challenge to the orthodox management of wounds. 2011. http://www.jameslindlibrary.org/articles/bartolome-hidalgo-de-agueros-16th-century-evidence-based-challenge-to-the-orthodox-management-of-wounds/. Accessed 9 Dec 2018.

10. 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.
11. Bilguer M. A dissertation on the inutility of the amputation of the limbs. London: Baldwin, Becket and de Hondt; 1764.
12. Alanson E. Practical observations on amputation, and the after-treatment. 2nd ed. London: Joseph Johnson; 1782.
13. Poisson D, Larrey D, editors. Recherches de statistique sur l'affection calculeuse, par M. le docteur Civiale. Comptes rendus hebdomadaires des séances de l'Académie de Sciences [Statistical research on (urinary) stone, by M. Dr Civiale]. Paris: Bachelier; 1835.
14. Cheseldon W. The anatomy of the human body. 5th ed. London: William Bowyer; 1740. http://www.jameslindlibrary.org/cheselden-w-1740/. Accessed 14 Dec 2018
15. Tröhler U. "To improve the evidence of medicine": the 18th century British origins of a critical approach. Edinburgh: Royal College of Physicians of Edinburgh; 2000.
16. Morabia A. Pierre-Charles-Alexandre Louis and the evaluation of bloodletting. J R Soc Med. 2006;99(3):158–60.
17. Ackerknecht EH. Medicine at the Paris hospital, 1794–1848. Baltimore: Johns Hopkins Press; 1967.
18. Tröhler U. Statistics and the British controversy about the effects of Joseph Lister's system of antisepsis for surgery, 1867–1890. 2014. http://www.jameslindlibrary.org/articles/statistics-and-the-british-controversy-about-the-effects-of-joseph-listers-system-of-antisepsis-for-surgery-1867-1890/. Accessed 9 Dec 2018.
19. Pernick M. A Calculus of suffering: pain, profession, and anesthesia in 19th century America. New York: Columbia University Press; 1985.
20. Chalmers I. Why the 1948 MRC trial of streptomycin used treatment allocation based on random numbers. J R Soc Med. 2011;104(9):383–6.
21. Hamilton AL. Dissertatio Medica Inauguralis De Synocho Castrensi. Edinburgh: J Ballantyne; 1816. Accessed 9 Dec 2018.
22. Hill B. Clinical lectures on the stricture of the urethra. BMJ. 1879;2:883–6.
23. Fox LW. Immediate capsulotomy following the removal of cataract. JAMA. 1894;22:837–9.
24. Langton J. A discussion on the treatment of hernia in children. BMJ. 1899;2:470–2.
25. Jones DS. Surgery and clinical trials: the history and controversies of surgical evidence. In: Schlich T, editor. The Palgrave handbook of the history of surgery. Montreal: Macmillan Publishers; 2018. p. 479–501.
26. Barnes BA. Discarded operations: surgical innovation by trial and error. Costs, risks and benefits of surgery. New York: Oxford University Press; 1977. p. 109–23.
27. Lerner BH. The breast cancer wars: fear, Hope, and the pursuit of a cure in twentieth-century America. New York: Oxford University Press; 2011.
28. Wilde S. See one, do one, modify one: prostate surgery in the 1930s. Med Hist. 2004;48:351–66.
29. HIlts P. Protecting America's health: the FDA, business, and one hundred years of regulation. Chapel Hill: University of North Carolina Press; 2003.
30. Amberson J, McMahon B, Pinner M. A clinical trial of sanocrysin in pulmonary tuberculosis. Am Rev Tuberc. 1931;24:401–35.
31. Brock B. The sanocrysin treatment of pulmonary tuberculosis in the white and negro races. Am Rev Tuberc. 1931;24:436–45.
32. Diaz M. Lessons from using randomization to assess gold treatment for tuberculosis. 2004. http://www.jameslindlibrary.org/articles/lessons-from-using-randomization-to-assess-gold-treatment-for-tuberculosis/. Accessed 10 Dec 2018.
33. Marshall G, Crofton J, Hurford J, Brent R, Ross I, Gilmour WS, et al. Streptomycin treatment of pulmonary tuberculosis: a medical research council investigation. Br Med J. 1948;2(4582):769–82.
34. Knapp MR, Beecher HK. Postanesthetic nausea, vomiting, and retching; evaluation of the antiemetic drugs dimenhydrinate (dramamine), chlorpromazine, and pentobarbital sodium. J Am Med Assoc. 1956;160(5):376–85.

35. Conn HO, Lindenmuth WW. Prophylactic Portacaval anastomosis in cirrhotic patients with esophageal varices—a preliminary report of a controlled study. N Engl J Med. 1962;266:743–9.
36. Brinkley D, Haybittle JL. Treatment of stage-II carcinoma of the female breast. Lancet. 1966;288:291–5.
37. Cobb L, Thomas G, Dillard D, Merendino K, Bruce R. An evaluation of internal-mammary-artery ligation by a double-blind technic. N Engl J Med. 1959;260:1115–8.
38. Crenner C. Placebos and the progress of surgery. In: Schlich T, Crenner C, editors. Technological change in modern surgery. Rochester: University of Rochester Press; 2017. p. 156–84.
39. Goligher JC, Pulvertaft CN, Watkinson G. Controlled trial of vagotomy and gastro- enterostomy, vagotomy and antrectomy, and subtotal gastrectomy in elective treatment of duodenal ulcer: interim report. Br Med J. 1964;1(5381):455–60.
40. Handley WS. Cancer of the breast and its operative treatment. London: Murray; 1906.
41. Ellis H. A history of surgery. London: Greenwich Medical Media; 2001.
42. Halsted WS. The results of radical operations for the cure of carcinoma of the breast. Ann Surg. 1907;46:1.
43. Lane-Claypon J. Reports on public health, medical subjects. London, H.M. Stationary Office; 1924.
44. William Stewart Halsted, surgical papers. Wellcome Collection. https://wellcomecollection.org/works/bpn6hxgq?query=L0004968.
45. McWhirter R. Simple mastectomy and radiotherapy in treatment of breast cancer. Brit J Radiol. 1955;28:128–39.
46. Williams I, Murley R, Curwen M. Carcinoma of the female breast: conservative and radical surgery. Brit Med J. 1953;2:787–96.
47. Mustakallio S. Treatment of breast cancer by tumor extirpation and roentgen therapy instead of radical operation. J Facult Radiol. 1956;6:23–6.
48. George Crile J. Simplified treatment of cancer of the breast: early results of a clinical study. Ann Surg. 1961;153:745–61.
49. Lee BJ. The therapeutic value of irradiation in the treatment of mammary cancer: a survey of five-year results in 355 cases treated at the memorial hospital of New York. Ann Surg. 1928;88(1):25–47.
50. Grace EJ. Simple mastectomy in cancer of the breast. Am J Surg. 1937;35:512–4.
51. Holman E. Sir William Osler, William Stewart Halsted, Harvey Cushing: some personal reminiscences. Surgery. 1965;57:589–601.
52. Kaae S, Johansen H. Simple mastectomy plus postoperative irradiation by the method of McWhirter for mammary carcinoma. Ann Surg. 1969;170(6):895–9.
53. Fisher B, Jeong J, Anderson S, Bryant J, Fisher E, Wolmark N. Twenty-five-year follow-up of a randomized trial comparing radical mastectomy, total mastectomy, and total mastectomy followed by irradiation. N Engl J Med. 2002;347(8):567–75.
54. McPherson K, Fox M. Treatment of breast cancer. In: Bunker J, Barnes B, Mosteller F, editors. Costs, risks and benefits of surgery. Oxford: Oxford University Press; 1977. p. 308–22.
55. Ellis J, Mulligan I, Rowe J, Sackett DL. Inpatient general medicine is evidence based. A-Team, Nuffield Department of Clinical Medicine. Lancet. 1995;346(8972):407–10.
56. Howes N, Chagla L, Thorpe M, McCulloch P. Surgical practice is evidence based. BJS. 1997;84(9):1220–3.
57. Kenny SE, Shankar KR, Rintala R, Lamont GL, Lloyd DA. Evidence-based surgery: interventions in a regional paediatric surgical unit. Arch Dis Child. 1997;76(1):50.
58. Wente MN, Seiler CM, Uhl W, Buchler MW. Perspectives of evidence-based surgery. Dig Surg. 2003;20(4):263–9.
59. Gelijns AC, Ascheim DD, Parides MK, Kent KC, Moskowitz AJ. Randomized trials in surgery. Surgery. 2009;145(6):581–7.
60. Horton R. Surgical research or comic opera: questions, but few answers. Lancet. 1996;347(9007):984–5.

61. Wilde S, Hirst G. Learning from mistakes: early twentieth-century surgical practice. J Hist Med Allied Sci. 2009;64(1):38–77.
62. McLeod RS. Issues in surgical randomized controlled trials. World J Surg. 1999;23:1210–4.
63. Reitsma AM, Moreno JD. Ethical regulations for innovative surgery: the last frontier? J Am Coll Surg. 2002;194(6):792–801.
64. McCulloch P, Altman D, Campbell WB, Flum D, Glasziou P, Marshall J, et al. No surgical innovation without evaluation: the IDEAL recommendations. Lancet. 2009;374:1105–12.
65. Love J. In comment: drugs and operations: some important differences. JAMA. 1975;232:37–8.
66. Macklin R. The ethical problems with sham surgery in clinical research. N Engl J Med. 1999;341(13):992.
67. Lederer SE. Subjected to science: human experimentation in American before the second world war. Baltimore: The Johns Hopkins University Press; 1995.
68. Cutter L. Walter reed, yellow fever, and informed consent. Mil Med. 2016;181(1):90–1.
69. Moreno JD. Undue risk: secret state experiments on humans. New York: W.H. Freeman; 2000.
70. Krantz J. New drugs and the Kefauver-Harris amendment. J New Drugs. 1966;6(22):77–9.
71. Jones D, Grady C, Lederer S. "Ethics and clinical research"—the 50th anniversary of Beecher's bombshell. N Engl J Med. 2016;374:2393–8.
72. 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.
73. Reverby S. Examining Tuskegee: the infamous syphilis study and its legacy. Chapel Hill: University of North Carolina Press; 2009.
74. OHRP. The Belmont report [updated 15 Mar 2016]. https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html. Accessed 16 Dec 2018.
75. Burger ISJ, Goodman SN. Ethical issues in evidence-based surgery. Surg Clin N Am. 2006;86:151–68.
76. ACo S. Statement on issues to be considered before new surgical technology is applied to the care of patients. Bull Am Coll Surg. 1995;80:46–7.
77. Rubin E, Lazar D. Clinical trials office: what is new in research administration? Washington, DC: Association for Academic Health Centers; 2009.
78. EBCTCG. Treatment of early breast cancer. Vol 1. Worldwide evidence, 1985–1990. Oxford: Oxford University Press; 1990. http://www.jameslindlibrary.org/early-breast-cancer-trialists-collaborative-group-1990/.
79. Kelahan A, Catalano R, Marinucci D. The history, structure, and achievements of the cancer cooperative groups. Managed Care Cancer. 2001:28–33.
80. Volpe M, Tocci G. 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.

Chapter 2
Ethics (Informed Consent and Conflicts of Interest)

Kara K. Rossfeld, Jordan M. Cloyd, Elizabeth Palmer, and Timothy M. Pawlik

2.1 Introduction

As a direct result of their health conditions, patients are inherently vulnerable in their relationship with physicians. Lacking the experience or knowledge required to understand complex medical decision, they must trust that their providers' recommendations are consistent with their values, goals, and expectations. Concurrently, physician-investigators witness gaps in clinical knowledge and are motivated to address these gaps through scientific discovery and clinical trials. The challenge inherent in this task is that patients enrolled in clinical research may or may not directly benefit from the research and its associated risks, although the welfare of future patients is dependent on this research. Maintaining the highest ethical standards is therefore not only important to protect the health and rights of patients involved in clinical trials but also to protect the integrity of the clinical research enterprise.

2.2 History of Modern Biomedical Ethics

Evidence of principles guiding ethical behavior in medicine have been found dating back to antiquity. These principles were evolving as the medical field advanced in the nineteenth and twentieth centuries, but the watershed moment that shaped current biomedical ethics was the 1947 Nuremberg trials of Nazi physicians who per-

K. K. Rossfeld (✉) · J. M. Cloyd · E. Palmer · T. M. Pawlik
The Ohio State University Wexner Medical Center, Columbus, OH, USA
e-mail: Kara.Rossfeld@osumc.edu

T. M. Pawlik, J. A. Sosa (eds.), *Clinical Trials*, Success in Academic Surgery,
https://doi.org/10.1007/978-3-030-35488-6_2

formed experiments on Jewish and other marginalized persons imprisoned in concentration camps. These trials elevated the public discourse and highlighted the need for an international consensus guiding human subjects research which led to the creation of the Nuremberg Code [1]. This laid the foundation for the Declaration of Helsinki created in 1964 by the World Medical Association [2]. The Declaration of Helsinki established several important principles:

- There must be a scientific basis of the proposed research (preceded by appropriate laboratory modeling).
- Risks to the research subjects must be carefully weighed against and must not exceed potential benefits to the subjects or society.
- Research may not be performed without the informed consent provided freely by the research subjects or their legal guardian.
- A research subject must be allowed to withdraw from a study at any time.

The Declaration of Helsinki has been revised and updated many times since the original drafting, including in 1975 at which time the World Medical Association recommended the establishment of independent committees to review and oversee proposed research protocols [3].

At the same time in the United States, there was public outcry for protections of human research subjects when it came to light in 1972 that hundreds of African American men with syphilis were being studied by the Public Health Service while withholding effective treatment and without consent [4]. Public outrage over the Tuskegee Study scandal as well as other growing concerns related to biomedical and pharmaceutical research [5] led to the creation of the National Research Act in 1974.

The National Research Act authorized the creation of the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. This body convened to establish detailed guidelines for human subjects research and also published the Belmont report, which established autonomy, beneficence, and justice as the primary principles guiding ethical clinical research. The National Commission laid the groundwork for legislation now referred to as the Common Rule, or 45 CFR 46 [6]. Adopted by 16 government agencies in 1991, this legislation formalized regulations for human subjects research. Among its many regulations was the requirement for institutional review boards (IRBs) to oversee publicly funded research in the United States.

2.2.1 The Final Rule

The Common Rule was revised in 2018 in an effort to modernize and reflect some of the current challenges in medical research. Debate began in 2015 with the release of the Notice of Proposed Rulemaking. Various stakeholders in healthcare,

biomedical research, pharmaceutical industry, and patient advocacy groups provided input to the proposed changes. The resulting changes accurately reflected the current state of biomedical ethics in the United States, capturing both the strong desire to preserve autonomy but also recognizing the social good that comes from biomedical research and the benefits of broad participation. For example, one proposed change was to prohibit the use of leftover deidentified biospecimens unless consent was obtained. This proposal was rejected by the public as being overly burdensome to researchers without delivering clear benefits or protections to patients [7].

The revised Common Rule, or Final Rule, was notable for several distinct changes. In an attempt to restore autonomy and self-determination to research subjects, informed consent documents must possess certain elements in an effort to make them more understandable for the lay public [8]. In addition, the federal code now allows patients to consent for future unspecified research on biospecimens, known as broad consent. The legislation decreases the burden on IRBs by expanding what research is included in the exempted category and decreasing review requirements of low-risk research, with the goal of increasing time for IRBs to attend to higher risk research. These changes to protections in federally funded research are required by law in an effort to regulate research integrity, but meeting these regulations alone does not inherently make endeavors in clinic research ethical [9].

2.3 The Seven Features of an Ethical Clinical Trial

While the aforementioned documents and other guidelines provide principles and regulatory measures for ethical human subjects research (Table 2.1), they fall short of providing a digestible summation of what makes a clinical trial ethical. In 2000, Emmanuel et al. sought to create an ethical framework by which to assess a proposed clinical trial for ethical shortcomings. Seven components are outlined that should be considered [10].

2.3.1 Societal Value

Nazi physicians put human subjects through horrific experiments in order to observe morbidity and mortality without any future societal value [11]. Not to be repeated, human subjects research must be of clear benefit to the field of medicine for the improvement of care of patients and society. There must be sufficient evidence to support the study, and there must be significant improvement in care anticipated for future patients.

Table 2.1 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.proxy.lib.ohio-state.edu/ohrp/archive/nurcode.html
Declaration of Helsinki	World Medical Association	1964	http://www.wma.net.proxy.lib.ohio-state.edu/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.proxy.lib.ohio-state.edu/ohrp/humansubjects/guidance/belmont.html
45 CFR 46 (Common rule)	US Department of Health and Human Services	1991	http://www.hhs.gov.proxy.lib.ohio-state.edu/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.proxy.lib.ohio-state.edu/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

2.3.2 Scientific Validity

The study must be scientifically sound. For trials with randomization, clinical equipoise must exist. Equipoise is the concept that there is genuine uncertainty as to the equivalency or superiority of one intervention compared to another [12]. If one therapy was thought to be superior, then patients in the other arm of the study would be receiving inferior care. This applies to procedural techniques as well. For example, endoscopic mucosal resection of an early esophageal cancer offers a much less morbid operation than esophagectomy. Data in patients with early esophageal cancers who were not surgical candidates first suggested that oncologic outcomes might be equivalent with an endoscopic resection while sparing morbidity of an esophagectomy [13, 14]. Once this was established [15], so too was the clinical equipoise to allow ethical evaluation in surgically fit patients who would otherwise be a candidate for an esophagectomy.

A flawed design or underpowered study is also ethically irresponsible. Not only do such studies pollute the literature and waste valuable resources, patients have been subject to the risks of the trial for no benefit [16]. Clinical researchers must have a firm understanding of how to formulate the right research questions, how to design appropriate trial protocols, and how to appropriately evaluate and interpret

data. Collaboration with statisticians is a moral imperative if the investigators leading a study do not have a deep understanding of study design and statistical analysis. Mentorship is invaluable while attaining these skills.

2.3.3 Fair Patient Recruitment

There must be justice in the method that participants are recruited to a study. Historically, research has often been performed on patients with insufficient socioeconomic means to pay for their care otherwise [17], violating the principles of justice and autonomy through coercion. For a study to be ethical, there must not be exploitation of vulnerable populations inherent in the population being studied. For this reason, incarcerated persons are often not included in human subjects research. On the other hand, there also must not be intentional exclusion of vulnerable populations without a sound scientific basis. The burden of research risk must be shared by all (not just economically disadvantaged patients), and there must not be special access to trials and treatments based on one's financial resources.

2.3.4 Satisfactory Balance of Risks and Benefits

Risks to patients enrolled in the study must be outweighed by potential benefits to them and/or to society at large. Preclinical laboratory and animal modeling should be performed in order to assess potential risks of new therapies or techniques. Whenever possible, exposure to risks should be minimized, appropriate monitoring of potential risks should be built into study design, and unanticipated harm to research participants must be reported to appropriate study personnel in a protocolized fashion.

2.3.5 Independent Oversight

There must be an oversight committee reviewing the proposed research, ensuring there are no violations of research ethics. For federally funded research, IRBs act as a check on the researchers both in meeting regulatory requirements for appropriate conduct of the study and also to identify potential conflicts of interest.

2.3.6 Informed Consent

To preserve autonomy, patients must be provided with the tools necessary to make a well-informed decision about whether to participate in a trial. With the revisions to the Common Rule, informed consent must begin with a concise summary of

information that would broadly assist with a patient's decision to participate. It should include all necessary information that a "reasonable person" would want to know before deciding whether or not to participate, a legal albeit vague standard. This includes the reason for the study, potential benefits, anticipated risks, and alternatives to trial participation. This process requires verification that these concepts are understood by the research subject, even though, as will be discussed below, this confirmation is often overlooked.

2.3.7 Protections of Research Subjects

Participants in trials must be granted certain protections, including the right to terminate participation without repercussion. Protected health information and patient privacy must be safeguarded. When unforeseen risks become apparent, appropriate actions to protect the health of participants must be taken.

2.4 Innovation in the Design of Clinical Trials

Defining clear benefits to patients and/or society can occasionally pose challenges in clinical trial design. This is particularly relevant in cancer research, as research and development of treatments and techniques is an inherently slow process. Significant time is required to take an idea through basic science and preclinical modeling, institutional review board approval, accrual of trial participants, implementation of the intervention, and observation of the outcomes. Meanwhile, breakthroughs may happen within a discipline that render the benefit of a particular study unclear or potentially even irrelevant.

There is therefore a moral imperative to ensure trial design is as efficient as possible. One brilliant example of this is the I-SPY 2 Trial: Investigation of Serial Studies to Predict Your Therapeutic Response with Imaging and Molecular Analysis 2 [18]. This adaptive trial has enabled the study of numerous molecular targeted therapies for breast cancer in the neoadjuvant setting in a relatively short period of time. Tumor biomarkers and molecular subtypes are factored to optimize the arm to which a patient is randomized. Response is determined by evaluating pathologic response at the time of surgery rather than longer endpoints such as overall survival. Effective therapies are fast-tracked to the market, ineffective therapies are dropped from the protocol, and new therapies enter the pipeline. Granted, not all cancers are amenable to this type of trial, but it is a prime example of a trial design that is flexible, responsive to new therapies entering the market, and efficient thereby maximizing potential benefits of the study.

2.5 Informed Consent for Clinical Trials

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. 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.

2.5.1 Patient Understanding Is Insufficient

While obtaining informed consent for a clinical trial is both ethically and legally required, research suggests that patients frequently enroll in trials without a firm understanding of what the trial entails, sometimes misunderstanding key concepts [19]. Some concepts that are second nature in the conduct of high-quality research like equipoise and randomization are particularly difficult for the general public to grasp [20]. One systematic review demonstrated that patients adequately under-

stood the aims of the trials in which they were enrolled in only 54% of studies [21]. Patients are motivated to join a study for a number of reasons, and their decision is often based on emotion—whether that is fear of not having an alternative or a desire to be altruistic while potentially being helped [22]. Many patients just assume that a new treatment is inherently better than the standard of care [23]. Regardless of motivation, a research subject may consent without understanding, but it must be recognized that the fault is not with the research subject. The onus is on the principal investigator and the research team to ensure patients understand what they are enrolling in.

2.5.2 Assessing Patient Understanding

While the revisions to the Common Rule represent a starting point to improve patient understanding from the informed consent process (Table 2.2), these rules are still fairly ambiguous. How to present information in a way that facilitates comprehension for participants is poorly understood. While there is a growing body of literature in this field, several obstacles exist, including determining whether comprehension should be evaluated subjectively, objectively, or based on patient perception of or satisfaction with their own understanding [24]. These methods all carry their own ethical limitations. The current practice of subjective assessment by the research team based on the general interactions during the informed consent process is inadequate—research subjects understand less than we think, violating their right to self-determination. An objective assessment of understanding (i.e., requiring potential research subjects to pass a quiz in order to enroll) on the other hand might also be problematic both from ethical and scientific standpoints as it may discourage or exclude patients with poor

Table 2.2 Requirements of informed consent

Document organization	Begins with concise summary of information
Required content	• Foreseeable risk or discomfort • Expected benefit to society or individual • Alternative treatments • Plan for maintaining confidentiality • Description of compensation for injury and explanation of medical therapies for said injury • Contact information for questions and to report concerns Right to discontinue participation
Excluded content	• Exculpatory language absolving research team from liability • Lists of facts

Adapted from Title 45: Public Welfare; Code of Federal Regulations Part 46.116 General requirements for Informed Consent

literacy from participating. This would violate the principle of justice and may lead to selection bias in the participant pool. Tools have been developed in an attempt to measure understanding, deliberation, and the informed consent process [25–28]. One of these tools measures the quality of the content delivered by the research coordinator and evidence of patient understanding during the consent process [29]. While this tool does not directly measure patient understanding, it may be useful in auditing the quality of the interactions to providing formative feedback to recruiters as a method for improving the informed consent process.

2.5.3 What Patients Want from the Informed Consent Process

What do "reasonable persons" desire from the informed consent process? One qualitative study demonstrated that patients want advanced warning that they will be approached for research [19]. Communication should be clear and devoid of jargon, and they want adequate time with their physician and research coordinator. It must be recognized that the amount of information and detail that patients want prior to making a decision about participation varies, and (as much as possible), the research team should respect these preferences [30]. As the task of deciding whether to participate in a trial can be daunting to patients, they should be given adequate time to consider their decision.

2.5.4 Optimal Presentation of Information

Studies have been performed to examine the impact of using audiovisual materials to enhance understanding, although many of these studies are methodologically limited. In general, there is not enough evidence to indicate that audiovisual materials enhance understanding of trials by research participants [31]. In one study comparing 18 research protocols presented to patients either via the traditional written format, a booklet, a computer-assisted instructional program, or an instructional video, no format proved superior in conveying information for patient understanding as judged by a knowledge quiz. Unfortunately, on average, no intervention group or demographic scored higher than 67% on the quiz [32].

While additional research is needed, digital platforms are a natural choice to present complex information in a simplified understandable way, supplemented with animations and graphics. This format carries potential for real-time assessment of understanding, with additional built-in educational modules tailored to concepts that patients incompletely understand [33]. Of course, this may be costly to develop, and further research is needed to know if this type of interactive platform would be helpful for increasing comprehension and/or recruitment.

2.5.5 Special Considerations for Informed Consent

Informed consent must be obtained without coercion to be ethical and valid. As such, the setting in which is obtained must be carefully considered. While any patient may be subject to coercion by their physician or research team, one must be particularly mindful of populations who may be vulnerable to agreeing to a trial without due consideration. Vulnerable populations should be protected by specific measures 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 whose voluntary written informed consents 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*.

Similarly, research in the elderly and the young deserves special mention. 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, an assent from the child should be obtained in conjunction with an appropriate consent from the parents [34]. Likewise, dissent from the child should also be appropriately respected. In research of the elderly, additional attention must be paid to ensure these patients have capacity for complex decision-making [35]. In situations where patients don't have capacity (e.g., in mild dementia), their surrogate decision-makers must be included in the consent process while also ideally securing assent to trial participation from the patient.

Research in the emergency setting is another situation where obtaining informed consent is difficult if not impossible to do. Patients may not be capable of providing consent due to the acuity of their condition. The shortened timeframe before intervention is also prohibitive for explaining the complexity of a trial and does not afford patients time to deliberate upon their decision. Regardless, patients with emergent surgical disease deserve high-quality research to inform their treatment. Regulatory bodies recognize this need and provide exceptions for obtaining informed consent for research on life-threatening conditions. For those conditions that are not immediately life-threatening, there are several solutions for trials in the emergency setting, first and foremost by engaging the community in which the research is being performed, educating the community and seeking stakeholder input [36]. 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 [37].

Lastly, clinical trials involving surgical innovation deserve mention. Surgical innovation has been responsible for some of the greatest revolutions in the field of medicine. Problem-solving is inherent in the work surgeons do, finding a way to accomplish the goal of an operation, developing tools to accomplish certain tasks, or decreasing morbidity. Often, there is not enough experience with new techniques or tools to know the actual level of risk for a known complication, especially when rates of the complication are low to begin with [23]. When enlisting patients in trials examining new techniques or tools, physician-investigators must be transparent with regard to the unknowns. If a surgeon is unable to be forthright about his or her lack of experience with a technique or therapy, it should not be offered [38].

2.6 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 (COI) may arise in the clinical research setting are the physician-researcher conflicts and financial conflicts. It is important to distinguish that a COI is the existence of competing interests which *may* lead to bias or result in harm. COI is often misconstrued as the occurrence of bias and harm [39]. There is a scale on which a secondary interest poses a risk of influence on a physician-investigator, and there is a range of harm that may be done as a result of perpetration of a COI [40]. For example, the risk of influence on a physician-investigator is less for accepting a free meal from a pharmaceutical company as compared to the risk of influence on a physician-investigator given stock options for a drug he or she is investigating. The range of harm done similarly can be broad, from undermining the patient-physician relationship by putting pressure on a patient to participate in a trial to gross underreporting of adverse effects or risks of an intervention. Many medical professional and industry organizations have published codes for ethical behavior related to managing conflicts of interests (Table 2.3), and researchers should be familiar with their own institutional regulations.

2.6.1 Physician-Researcher Conflicts

Perhaps the most subtle form of COI and the most difficult to detect occurs when physician-investigators become invested in a particular hypothesis or study. This could be because a physician becomes convinced of the benefit of an intervention from personal experience, despite equipoise in the larger community. In these

Table 2.3 Guidelines for managing conflict of interest

Organization	Document	Available at
American College of Physicians	"Conflicts of Interest" ACP Ethics Manual, 7th Edition	https://www.aamc.org/download/482216/data/protectingpatients.pdf
American College of Surgeons	Code of Professional Conduct	https://www.facs.org/about-acs/statements/stonprin
American Medical Association	Conflicts of Interest in Research: Code of Medical Ethics Opinion 7.1.4	https://www.ama-assn.org/delivering-care/ethics/conflicts-interest-research
Association of American Medical Colleges	Protecting Patients, Preserving Integrity, Advancing Health: Accelerating the Implementation of COI Policies in Human Subjects Research	https://www.aamc.org/download/482216/data/protectingpatients.pdf
Advanced Medical Technology Association	AdvaMed Code of Ethics on Interactions with Health Care Professionals	https://www.advamed.org/resource-center/advamed-code-ethics-interactions-health-care-professionals
Pharmaceuticals and Research and Manufacturers of America (PhRMA)	PhRMA Principles on Conduct of Clinical Trials	https://www.phrma.org/codes-and-guidelines/phrma-principles-on-conduct-of-clinical-trials

instances, there may be bias of the information the physician-investigator presents to possible research participants, including the overselling of potential benefits of joining a study or minimizing potential risks and options for alternative therapies, undermining the informed consent process and the patient's autonomy.

More visible yet are the tremendous pressures to produce results and publish, including for promotion and tenure, in securing research funding (an increasingly competitive resource), and in the cultivation of one's local and national reputation. These pressures may bias the physician-investigator in any number of ways, including in how he or she recruits patients, manages or interprets data, or publishes research findings.

2.6.2 Financial COI

Financial interests are the most visible forms of COI. Collaboration between physician-investigators and industry is critical in bringing forward new therapies and technologies to benefit the care of patients, but these relationships must be managed with caution [41]. The late 2000s saw a number of lawsuits against pharmacologic companies for illegal marketing practices and inappropriate payments to healthcare providers. As a part of the Affordable Care Act in 2010, the Physician Payments Sunshine Act was enacted requiring disclosure of industry payments to

physicians in an effort to increase transparency [42]. While it represents a step forward in accountability, there is still room for improvement in disclosure of COI in medical literature [43]. Disclosure requires self-regulation that is inherently limited by personal biases and self-perceptions and often disincentivizes physicians from reporting by requiring additional paperwork or administrative procedures for the evaluation of disclosed interests. Furthermore, there must be appropriate evaluation of disclosed interests to determine if further actions must be taken by the institution, and this may not occur effectively or at all [44].

Monetary reward for innovation, ingenuity, and improvement in most other industries would go unquestioned. There are no easy solutions to the ethical dilemmas inherent in how physicians are rewarded to do the same [45], but we must continue to examine these challenges, seek solutions, and strive to maintain the integrity of the research we perform, so that we may be worthy to serve the suffering through clinical care and research.

2.7 Conclusions

The medical field continues to make impressive strides forward in research and development. As it does so, there will continue to be challenges in the ethical nature of how science is applied. As demonstrated, the challenges to be met are multifaceted from understanding how we execute the very principles that have dominated biomedical research ethics such as informed consent to how we might revolutionize research to make it more efficient and worthwhile. It is critical to public trust and the special privilege afforded to physician-investigators that we are able to navigate ethical dilemmas, recognize when ethical lines are being crossed, and effectively self-regulate. It is equally important to recognize that the research field will present unforeseen ethical dilemmas, just as breakthroughs today were unimaginable at the turn of the last century. The time to hone these skills is now, so that we may be equipped for ethical challenges that lay ahead.

Acknowledgments We appreciate the outstanding work of Jukes Namm and Peter Angelos for authoring the first edition of this chapter.

References

1. Trials of war criminals before the Nuremberg Military tribunals under control council law no. 10. Washington, DC: US Government Printing Office; 1949. p. 181–182.
2. Declaration of Helsinki: recommendations guiding doctors in clinical research. Finland: 18th World Medical Assembly; June 1964.
3. Declaration of Helsinki: recommendations guiding medical doctors in biomedical research involving human subjects. Japan: 29th World Medical Assembly; October 1975.
4. Brandt AM. Racism and research: the case of the Tuskegee syphilis study. Hast Cent Rep. 1978;8(6):21–9.

5. Beecher HK. Ethics and clinical research. N Engl J Med. 1966;274(24):1354–60.
6. Sparks J. Timeline of laws related to the protection of human subjects. National Institutes of Health Office of History; June 2002. https://history.nih.gov/about/timelines_laws_human.html.
7. Menikoff J. Rules and regulations. Fed Regist. 2017;82(12):7149–62.
8. Menikoff J, Kaneshiro J, Pritchard I. The common rule, updated. N Engl J Med. 2017;376(7):613–5.
9. Electronic Code of Federal Regulations. In: Services UDoHaH, editor. Title 45: public welfare; part 46—protection of human subjects. Online 2018.
10. Emanuel EJ, Wendler D, Grady C. What makes clinical research ethical? JAMA. 2000;283(20):2701–11.
11. Barondess JA. Medicine against society. Lessons from the Third Reich. JAMA. 1996;276(20):1657–61.
12. Freedman B. Equipoise and the ethics of clinical research. N Engl J Med. 1987;317(3):141–5.
13. Berry MF, Zeyer-Brunner J, Castleberry AW, Martin JT, Gloor B, Pietrobon R, et al. Treatment modalities for T1N0 esophageal cancers: a comparative analysis of local therapy versus surgical resection. J Thorac Oncol. 2013;8(6):796–802.
14. Ell C, May A, Gossner L, Pech O, Gunter E, Mayer G, et al. Endoscopic mucosal resection of early cancer and high-grade dysplasia in Barrett's esophagus. Gastroenterology. 2000;118(4):670–7.
15. Omae M, Fujisaki J, Horiuchi Y, Yoshizawa N, Matsuo Y, Kubota M, et al. Safety, efficacy, and long-term outcomes for endoscopic submucosal dissection of early esophagogastric junction cancer. Gastric Cancer. 2013;16(2):147–54.
16. Ioannidis JPA, Stuart ME, Brownlee S, Strite SA. How to survive the medical misinformation mess. Eur J Clin Investig. 2017;47(11):795–802.
17. Rice TW. The historical, ethical, and legal background of human-subjects research. Respir Care. 2008;53(10):1325–9.
18. 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(1):97–100.
19. Behrendt C, Golz T, Roesler C, Bertz H, Wunsch A. What do our patients understand about their trial participation? Assessing patients' understanding of their informed consent consultation about randomised clinical trials. J Med Ethics. 2011;37(2):74–80.
20. Robinson EJ, Kerr CE, Stevens AJ, Lilford RJ, Braunholtz DA, Edwards SJ, et al. Lay public's understanding of equipoise and randomisation in randomised controlled trials. Health Technol Assess. 2005;9(8):1–192, iii-iv.
21. Falagas ME, Korbila IP, Giannopoulou KP, Kondilis BK, Peppas G. Informed consent: how much and what do patients understand? Am J Surg. 2009;198(3):420–35.
22. Dellson P, Nilsson K, Jernstrom H, Carlsson C. Patients' reasoning regarding the decision to participate in clinical cancer trials: an interview study. Trials. 2018;19(1):528.
23. Angelos P. Ethical issues of participant recruitment in surgical clinical trials. Ann Surg Oncol. 2013;20(10):3184–7.
24. Bossert S, Strech D. An integrated conceptual framework for evaluating and improving 'understanding' in informed consent. Trials. 2017;18(1):482.
25. Joffe S, Cook EF, Cleary PD, Clark JW, Weeks JC. Quality of informed consent: a new measure of understanding among research subjects. J Natl Cancer Inst. 2001;93(2):139–47.
26. Cohn EG, Jia H, Smith WC, Erwin K, Larson EL. Measuring the process and quality of informed consent for clinical research: development and testing. Oncol Nurs Forum. 2011;38(4):417–22.
27. Gillies K, Elwyn G, Cook J. Making a decision about trial participation: the feasibility of measuring deliberation during the informed consent process for clinical trials. Trials. 2014;15:307.

28. Stryker JE, Wray RJ, Emmons KM, Winer E, Demetri G. Understanding the decisions of cancer clinical trial participants to enter research studies: factors associated with informed consent, patient satisfaction, and decisional regret. Patient Educ Couns. 2006;63(1–2):104–9.
29. Wade J, Elliott D, Avery KNL, Gaunt D, Young GJ, Barnes R, et al. Informed consent in randomised controlled trials: development and preliminary evaluation of a measure of participatory and informed consent (PIC). Trials. 2017;18(1):327.
30. Childers R, Lipsett PA, Pawlik TM. Informed consent and the surgeon. J Am Coll Surg. 2009;208(4):627–34.
31. Synnot A, Ryan R, Prictor M, Fetherstonhaugh D, Parker B. Audio-visual presentation of information for informed consent for participation in clinical trials. Cochrane Database Syst Rev. 2014(5):CD003717.
32. Agre P, Rapkin B. Improving informed consent: a comparison of four consent tools. IRB. 2003;25(6):1–7.
33. Grady C, Cummings SR, Rowbotham MC, McConnell MV, Ashley EA, Kang G. Informed consent. N Engl J Med. 2017;376(9):856–67.
34. Wendler D. The assent requirement in pediatric research. In: Emanuel EJ, editor. New York: Oxford University Press; 2008.
35. Gilbert T, Bosquet A, Thomas-Anterion C, Bonnefoy M, Le Saux O. Assessing capacity to consent for research in cognitively impaired older patients. Clin Interv Aging. 2017;12:1553–63.
36. Karlawish J. Emergency research. In: Emanuel EJ, editor. New York: Oxford University Press; 2008.
37. Capron A. Legal and regulatory standards of informed consent in research. In: Emanuel EJ, editor. New York: Oxford University Press; 2008.
38. Angelos P. Surgical ethics and the challenge of surgical innovation. Am J Surg. 2014;208(6):881–5.
39. Stead WW. The complex and multifaceted aspects of conflicts of interest. JAMA. 2017;317(17):1765–7.
40. Thompson DF. Understanding financial conflicts of interest. N Engl J Med. 1993;329(8):573–6.
41. Lo B, Field MJ, editors. Conflict of interest in medical research, education, and practice. The National Academies Collection: Reports funded by National Institutes of Health. Washington, DC: National Academies Press; 2009.
42. Agrawal S, Brennan N, Budetti P. The sunshine act—effects on physicians. N Engl J Med. 2013;368(22):2054–7.
43. Cherla DV, Olavarria OA, Holihan JL, Viso CP, Hannon C, Kao LS, et al. Discordance of conflict of interest self-disclosure and the centers of Medicare and Medicaid services. J Surg Res. 2017;218:18–22.
44. Jacmon H. Disclosure is inadequate as a solution to managing conflicts of interest in human research. J Bioeth Inq. 2018;15(1):71–80.
45. Lichter AS. Conflict of interest and the integrity of the medical profession. JAMA. 2017;317(17):1725–6.

Chapter 3
Generating a Testable Hypothesis and Underlying Principles of Clinical Trials

Cecilia G. Ethun and Shishir K. Maithel

A well-designed and executed clinical trial can be one of the most powerful and definitive ways to assess the effectiveness and safety of an intervention(s). Thus, thorough knowledge of the underlying principles of clinical trials is required for any investigator choosing to embark on such an endeavor. This chapter will elaborate on three important aspects of clinical trial development: defining clinical trials, developing a research question, and generating a testable hypothesis.

3.1 Defining Clinical Trials

We define a clinical trial as a *prospective study that employs one or more interventions and evaluates the subsequent effect on one or more outcomes in human subjects.* Inherent in this definition are several key features that differentiate clinical trials from other types of research studies [1]. First, a clinical trial must be *prospective*, not retrospective. Study participants are followed forward from a prespecified and well-defined point in time, known as "time zero." How time zero is defined depends on the type of intervention, the outcome measure in question, and the trial design. Further discussion of trial design can be found in Chap. 4. In contrast to a clinical trial, a case-control study is a retrospective study that identifies subjects based on the presence or absence of a disease or outcome and looks backward in time to examine the subjects' exposure to particular risk factors.

A clinical trial must also involve at least one *intervention*. This may be a single intervention or a combination of interventions; may be diagnostic, preventative,

C. G. Ethun · S. K. Maithel (✉)
Division of Surgical Oncology, Department of Surgery, Winship Cancer Institute, Emory University, Atlanta, GA, USA
e-mail: smaithe@emory.edu

T. M. Pawlik, J. A. Sosa (eds.), *Clinical Trials*, Success in Academic Surgery,
https://doi.org/10.1007/978-3-030-35488-6_3

therapeutic, or educational in nature; and may utilize drugs, devices, techniques, systems, programs, and/or schedules. Regardless of how the intervention is structured, it should be predefined by investigators and applied to study subjects in a standardized manner with the intention of having some *effect* (or lack thereof) on the *outcome(s)*. Methods for ensuring quality and standardization are discussed in Chap. 7. A study that follows subjects prospectively but does not involve an active intervention is considered an observational study, not a clinical trial.

Finally, although similar principles can be employed in plant-, animal-, and laboratory-based studies, clinical trials must involve *human subjects* [2]. This important distinction requires careful consideration and adherence to ethical standards and safety guidelines by the investigators, which are discussed in detail in Chaps. 2, 11, and 12.

3.2 Developing a Research Question

Once an investigator understands the basics of what a clinical trial is, they can start to focus on the development of their own study, beginning with a question. After all, every good clinical trial starts with a good question. While not specific to clinical trials, using the FINER criteria can help guide investigators when thinking about the question that their clinical trial intends to address [3]. First, the question should be *feasible*. Feasibility considerations include the cost of the trial and available funding, the expected duration, the sample size needed, the difficulty of and expertise required for the intervention, and access to the patient population being studied. While the final details of the study are not necessary at this stage, performing crude calculations and having general knowledge of what your trial might entail are important aspects of determining trial feasibility when formulating your study question [4–6]. Second, the question must be *interesting*. That is not to say that it has to be headline grabbing, about an en vogue topic, or applicable to society as a whole. But it should be of interest to the investigators, the study subjects, and the intended audience, however broad or narrow. Third, the question should be *novel*. It should fill a gap in, expand upon, or refute the existing literature on the topic [4].

Fourth, the question should be *ethical*. The ethics of a study question involves both general research ethics and the concept of equipoise [7]. This principle relies on the investigators and expert community being relatively uncertain of the merits (or lack thereof) of the intervention(s) on the proposed outcome. One classic example is the parachute question: does using a parachute improve survival in skydivers jumping at 10,000 feet compared to not using a parachute? There is no uncertainty that using a parachute improves survival and, more importantly, that not using a parachute would result in certain death; thus, there cannot possibly be equipoise to justify such a study. The concept of equipoise is also one of the primary drivers for the relative decrease in the use of placebo or no treatment control arms. That is, in many cases, a well-established standard-of-care treatment already exists and is often based on the results of first generation trials that demonstrated efficacy of an inter-

vention over a placebo or no treatment. Therefore, in order to maintain equipoise, any future clinical trials should be designed to compare newer interventions to the existing standard-of-care. Fifth, the question should be *relevant*. It should be important and contribute both to current and future research, as well as to patient care.

While the FINER criteria can help investigators think about the study question in broad terms, using the PICO format helps investigators write out the specifics of the research question [4]. PICO stands for the patient population (P) of interest, the planned intervention (I), the comparison (C) or control group, and the outcome (O) being measured. Time (T) is often added to PICO to describe the time frame during which the study will take place. Using the PICO(T) format to develop the study question enables investigators to establish the framework of the study upfront, which can then be used as a guide for specific inclusion and exclusion criteria, choice of study design, and determining the outcome measurement tool [4, 5].

3.3 Generating a Testable Hypothesis

The next step for investigators is to frame the study question in the form of a testable hypothesis. There are several key points to keep in mind when generating a hypothesis. First, the hypothesis should be a statement, not a question. It should be clear and declarative, avoiding words like, "may," "might," or "could," and it should be specific [8]. If the investigators follow the PICO(T) format when developing their study question, being specific in writing the hypothesis—that is, including the patient population, the intervention, and the comparison or control group—should be easy [4].

The hypothesis should also include a prediction regarding the outcome, which may be directional or nondirectional [5]. In general, in clinical trials, there are three possible outcomes: that the intervention has a positive effect, has a negative effect, or has no effect. The directionality of a hypothesis refers to whether the investigators predict that the difference in the outcome will be in one specific direction (either explicitly positive or explicitly negative). An example of a directional hypothesis is that drug X will *improve* survival compared to drug Y. Stating that drug X will *affect* survival compared to drug Y is an example of a nondirectional hypothesis—investigators are merely predicting that drug X will change the outcome in some way, but don't explicitly state whether that will be positive or negative.

Before choosing whether or not to use a directional hypothesis, however, it is important for investigators to understand the statistical implications of that decision—that is, the directionality of a hypothesis directly affects how the null (H_0) and alternative hypotheses (H_A) are defined and can be used to justify either one-sided (for directional hypotheses) or two-sided (for nondirectional hypotheses) tests of significance [9]. Using a directional hypothesis, H_A may be that drug X improves survival, while H_0 is that drug X does not improve survival. Although this may seem reasonable, it is important to realize that hidden within H_0 are both the possibilities that drug X has no effect on survival or that drug X actually *worsens* survival. In this example, a one-sided test of significance would not differentiate between drug X

having no effect and having a negative effect on survival, even if the "truth" were that drug X worsens survival. While one-sided tests have greater statistical power to detect a difference in the predicted direction (i.e., it would be easier to find a significant improvement in survival if the "truth" was, in fact, that drug X improved survival), the possibility of failing to uncover the opposite outcome (in this case, that drug X is harmful) is enough for many investigators to avoid using one-sided tests. Indeed, two-sided tests of significance are more commonly used and most often preferred [5, 9].

Next, the hypothesis must be testable. The study variables must lend themselves to being observed, measured, and analyzed, and there has to be more than one possible outcome. To that end, the hypothesis cannot be an opinion or a fact. A testable hypothesis is also one that is feasible, a feature that has been discussed previously in this chapter with regard to the study question. Patient advocacy groups must be utilized and consulted when designing a study and generating a hypothesis as ultimately, a clinical trial will only be successful if patients agree to enroll. Finally, and perhaps most importantly, the hypothesis should be written a priori, or before the data are collected. The hypothesis should drive collection of the data, not be written as a result of the data.

3.4 Conclusion

Writing and executing a clinical trial is a daunting task, and one that is often unsuccessful for a wide variety of reasons. Because of this, understanding the foundational concepts of clinical trials—the definition of a clinical trial, the study question, and the hypothesis—is of critical importance. Mastery of these principles will give investigators a strong start in the right direction and the best opportunity for success.

References

1. NIH's definition of a clinical trial. National Institutes of Health: grants & funding. 2019. https://grants.nih.gov/policy/clinical-trials/definition.htm. Accessed 15 June 2019.
2. Definition of human subjects research. National Institutes of Health: grants & funding. 2019. https://grants.nih.gov/policy/humansubjects/research.htm. Accessed 15 June 2019.
3. Hulley SB, Cummings SR, Browner WS, Grady DG, Newman TB. Designing clinical research. 3rd ed. Philadelphia: Lippincott Williams and Wilkins; 2007.
4. Brian Haynes R. Forming research questions. J Clin Epidemiol. 2006;59(9):881–6.
5. Farrugia P, Petrisor BA, Farrokhyar F, Bhandari M. Practical tips for surgical research: research questions, hypotheses and objectives. Can J Surg. 2010;53(4):278–81.
6. Friedman LM, Furberg CD, DeMets DL, Reboussin DM, Granger CB. Fundamentals of clinical trials. 5th ed. New York, NY: Springer; 2015.
7. Freedman B. Equipoise and the ethics of clinical research. N Engl J Med. 1987;317(3):141–5.
8. Nottage M, Siu LL. Principles of clinical trial design. J Clin Oncol. 2002;20(18 Suppl):42S–6S.
9. Bland JM, Altman DG. One and two sided tests of significance. BMJ. 1994;309(6949):248.

Chapter 4
Trial Design: Overview of Study Designs

Puneet Singh, Yu Shen, and Kelly K. Hunt

4.1 Introduction

Clinical trials are fundamentally the investigation of human subjects under experimental conditions. For these trials to be successful, they must be well planned in the design phase to evaluate the prespecified outcomes. Thus, it is imperative for the investigators to understand clinical trial design which follows a typical progression from preclinical studies in animals to phase I–IV trials as described in this chapter (Fig. 4.1).

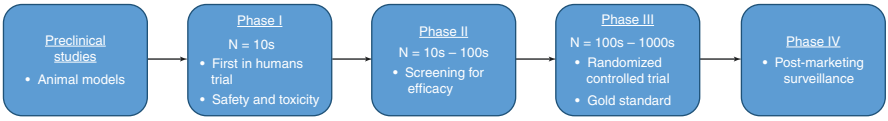


Fig. 4.1 Overview of clinical trials

P. Singh · K. K. Hunt (✉)
Division of Surgery, Department of Breast Surgical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA
e-mail: khunt@mdanderson.org

Y. Shen
Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

T. M. Pawlik, J. A. Sosa (eds.), *Clinical Trials*, Success in Academic Surgery,
https://doi.org/10.1007/978-3-030-35488-6_4

4.2 Phase I Trials

Once preclinical studies in animals are conducted to determine potentially appropriate doses in humans, a phase I trial is initiated as the "first-in-humans" trial. Phase I trials are also referred to as dose escalation or human pharmacology studies, and the goal is to test safety and toxicity of the experimental condition and identify a safe dose for humans, typically for a pharmaceutical drug [1]. Furthermore, the enrolled subjects do not have to have the same disease condition; for example, a drug or combination of drugs may be tested in patients with different solid tumors. The primary endpoint is determination of a safe dose which differs based on the type of drug. For cytotoxic medications, this dose is known as the maximal tolerated dose (MTD) in contrast to biomarker-based dose finding for targeted medications where the optimal and safe dose may be a better endpoint [2, 3]. Additionally, phase I trials may include secondary endpoints such as drug tolerability, dosing interval, route of delivery, and pharmacokinetics/pharmacodynamics (PK/PD) [4]. As this is the first phase in testing the safety of a drug, the number of subjects enrolled will be on the order of tens.

The most commonly utilized strategy for dose escalation, particularly for cytotoxic drugs, is the 3 + 3 design, even though the design has been found to have poor performance in general. In this dose escalation strategy, three subjects receive a medication at a dose based on preclinical studies and are monitored for dose-limiting toxicities (DLTs). Successive cohorts of three patients are enrolled with increasing doses until a DLT occurs. If one DLT occurs, then an additional three patients are given that dose and monitored; however, if two or three DLTs occur, then the next lower dosing level is expanded with three patients [4, 5]. This strategy is used until reaching the MTD which is subsequently the dose for phase II trials [5]. Due to the limitations and suboptimal performance of the 3 + 3 design, additional dose escalation strategies have emerged, including the accelerated-titration design and the model-based designs. The accelerated-titration design allows for rapid increases in the initial doses evaluated in single-patient cohorts until a DLT occurs at which point the strategy reverts to the traditional 3 + 3 design [5]. Model-based designs, of which the continual reassessment method based on Bayesian principles is one example, use data from all prior treated patients to mathematically model dose and toxicity [4–6]. These strategies may be more efficient in determining the MTD and allowing for a more rapid transition to phase II studies.

In oncology, there has been a significant growth in molecularly targeted agents (MTAs), and these agents are introduced in phase I studies. The PARP inhibitor, Olaparib, introduced in an enriched cohort of BRCA I and II mutation carriers with advanced solid tumors, is an example of this type of agent [7]. This trial used a modified accelerated-titration design with three patients enrolled at a dose that was doubled if no DLTs occurred versus expansion of the cohort to six if one DLT occurred. If two DLTs were observed, that was determined to be the maximum administered dose. Table 4.1 depicts the dose escalation strategy employed in the trial. The objectives were consistent with a phase I trial: evaluation of safety, docu-

Table 4.1 Dose-escalation scheme utilized in a phase I trial of a PARP inhibitor, Olaparib, in an enriched cohort of BRCA I and II mutation carriers with advanced solid tumors

Dose level and schedule	Number of patients	Total number of cycles	Median number of cycles	Number of dose-limiting toxicities in cycle 1 (%)
10 mg daily. 2 out of 3 weeks	3	12	2	0(0)
20 mg daily, 2 out of 3 weeks	3	5	2	0(0)
40 mg daily, 2 out of 3 weeks	5	15	2	0(0)
80 mg daily, 2 out of 3 weeks	3	9	4	0(0)
60 mg bid. 2 out of 3 weeks	4	10	2	0(0)
100 mg bid. 2 out of 3 weeks	4	15	3	0(0)
100 mg bid continuously	5	8	2	0(0)
200 mg bid continuously	20	94	4	0(0)
400 mg bid continuously	8	41	2.5	1(12.5)
600 mg bid continuously	5	13	2	2(40)

Reprinted with permission [7]

mentation of adverse events, DLTs, determination of the MTD, and the PK/PD profile of the drug. This trial identified the doses used for the subsequent phase II trials of Olaparib in advanced breast [8] and ovarian cancer [9].

4.3 Phase II Trials

The primary purpose of phase II trials is to test the early efficacy of a drug or intervention and may be referred to as "therapeutic exploratory" studies [1]. They serve as screening trials for effective drugs that can move on to pivotal phase III trials, which are more costly [10]. Conducted in a larger (on the order of tens to hundreds) cohort of patients, phase II trials are often single arm studies in one disease type that measure objective response as the primary endpoint of efficacy [11]. Other traditional endpoints of efficacy such as disease-free survival may be included. Occasionally, in phase IIB trials, there is a control arm with or without randomization or a comparison to a historical control group. The goal is to provide preliminary findings for hypothesis generation or designing future phase III trials, since phase II trials are often not adequately powered to show efficacy for the most important clinical endpoints of interest (e.g., disease-free or overall survival) [1, 5]. Another

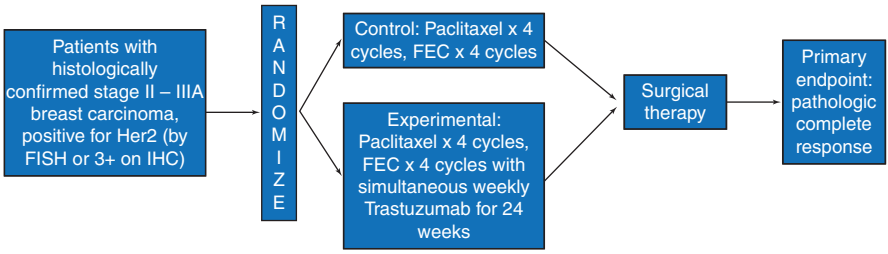


Fig. 4.2 ACOSOG Z1041, a phase II trial of neoadjuvant systemic therapy with standard chemotherapy and Trastuzumab. *FEC* fluorouracil, epirubicin, cyclophosphamide [12]

secondary objective of phase II trials is to further refine the safety profile of a drug including adverse events and PK/PD data [1, 5]. Phase II trials of medical devices similarly test efficacy and safety in a small cohort of subjects.

The early trials investigating trastuzumab (Herceptin®, Genentech Inc.) in human epidermal growth factor receptor 2 (HER2)-positive breast cancer are exemplary of the progression of drug development through the phases of clinical trials. Following on the finding that trastuzumab was more effective in treating HER2-positive metastatic breast cancer when given in combination with chemotherapy, Buzdar et al. designed a phase II study to assess the efficacy of this treatment strategy in the neoadjuvant (preoperative) setting in early stage, operable breast cancer patients. Patients were randomized to a chemotherapy regimen consisting of 4 cycles of paclitaxel followed by 4 cycles of 5-fluoruracil, epirubicin, and cyclophosphamide versus the same chemotherapy regimen with weekly trastuzumab for 24 weeks (Fig. 4.2) [12]. The primary endpoint was pathologic complete response (pCR; no evidence of invasive disease) in the breast and axilla. For all 42 randomized patients, the addition of trastuzumab resulted in a significantly higher pCR rate of 65.2% compared to 26.3% with chemotherapy alone ($p = 0.016$). This resulted in the institutional data monitoring committee recommending discontinuation of the control arm. Secondary safety endpoints included toxicity data and instances of dose reduction. The promising preliminary data on high pCR rates with the addition of trastuzumab to anthracycline- and taxane-based chemotherapy led to the development of the phase III ACOSOG Z1041 trial evaluating the impact of sequential versus concurrent delivery of trastuzumab with anthracyclines on pCR rates in HER2-positive breast cancer [13, 14].

4.4 Phase III Trials

Phase III trials or "therapeutic confirmatory" or "comparative efficacy" trials are the gold standard of evidence-based medicine. They are most often double-blinded, randomized controlled trials (RCTs) that evaluate efficacy and safety of a drug with a population of hundreds to thousands of patients and compare the drug, interven-

tion, or medical device to a control arm. Phase III trials are typically more costly and take longer to complete; thus, it is imperative to appropriately screen potential therapies in phase II studies. Factorial design, specifically the 2 × 2 design, is a methodology to evaluate two different drugs or interventions at the same time by randomizing patients to treatment A, treatment B, both, or none. Although this allows for simultaneous assessment of two experimental conditions, it assumes no interaction between the conditions which may be a limitation in the analysis and interpretation of this type of study [15]. Newer adaptive designs combine phase II (learning stage) with phase III (confirmatory stage) in a seamless phase II/III trial allowing for faster and less costly drug development [5, 16].

There are two main categories of phase III trials: comparative efficacy or equivalency trials. Comparative efficacy studies are the most common and are superiority trials that compare the experimental arm to a control arm, which may be standard of care or placebo, to determine if the experimental condition is superior. The objective of equivalency trials is to demonstrate that the experimental therapy is equivalent to the control within a prespecified margin. A subset of equivalency trials is the non-inferiority trial that investigates if the experimental arm is not less effective than the control arm again within a prespecified margin [1]. Interpretation of results differs based on the intention of the trial.

Several landmark RCTs were conducted by the National Surgical Adjuvant Breast and Bowel Project (NSABP) and randomized patients to different surgical treatments. NSABP B-04 was one of the earliest trials evaluating the local-regional management of breast cancer [17, 18]. Over 1000 clinically node-negative patients were randomized to radical mastectomy, total mastectomy with axillary lymph node dissection (ALND), or total mastectomy with regional irradiation and ALND only if the patient developed clinically positive nodes in follow-up. An additional 586 subjects who were clinically node positive were randomized to radical mastectomy or total mastectomy with regional irradiation (Fig. 4.3). Although survival rates were lower for patients with clinically positive nodes versus those with clinically

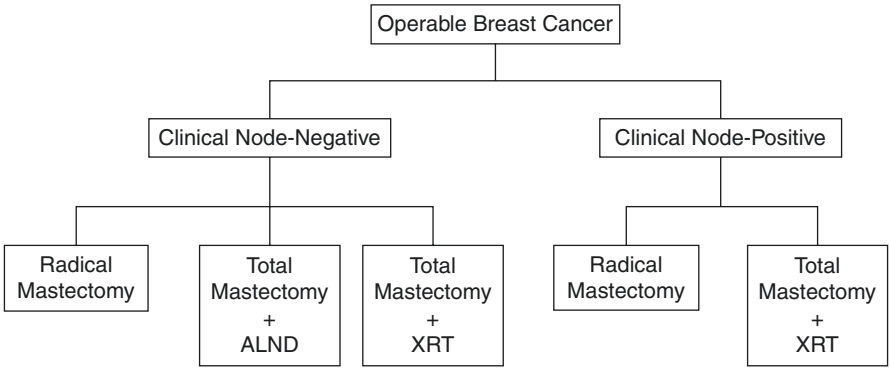


Fig. 4.3 NSABP B-04—Phase III trial design with randomization occurring to surgical treatments based on nodal status. *ALND* axillary lymph node dissection, *XRT* regional nodal irradiation. (Adapted with permission [18])

negative nodes, there was no difference in survival endpoints between the different treatment arms in each group. Thus, radical mastectomy, which had long been the standard of care for breast cancer patients, was replaced with less radical surgical approaches [17]. More recently, the ACOSOG Z0011 trial was designed as a phase III non-inferiority trial to evaluate the clinical outcomes of ALND versus no specific axillary treatment in early stage breast cancer patients with positive sentinel lymph nodes [19]. A total of 891 subjects with early stage, clinically node-negative breast cancer found to have 1–2 positive sentinel lymph nodes at the time of breast conservation surgery were randomized to ALND or no further axillary surgery (SLND alone) (Fig. 4.4). The primary endpoint was overall survival; the authors concluded

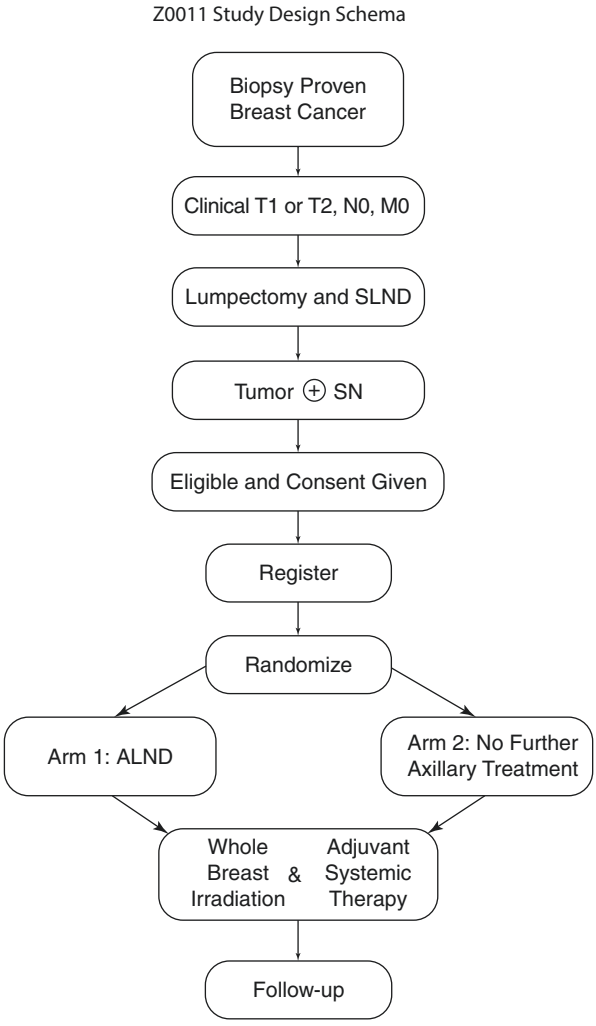


Fig. 4.4 ACOSOG Z0011 phase III trial design. *SLND* sentinel lymph node dissection, *ALND* axillary lymph node dissection. (Reprinted with permission [20])

that sentinel lymph node dissection alone did not result in inferior overall survival compared to ALND for this well-defined patient population [19, 20]. Both NSABP B-04 and ACOSOG Z0011 are examples of phase III trials that were designed with survival endpoints as the primary objective, as is common in oncology RCTs, and were practice changing.

The vast majority of phase III trials are pharmaceutical trials, whereas medical device and surgical intervention trials make up a smaller percentage [21, 22]. This is in part due to both real and perceived limitations and barriers that exist with non-pharmacologic trials. Surgical trials are more likely to be discontinued compared to nonsurgical trials due to poor recruitment of subjects. Lack of funding and negative interim results may further contribute to trial closure earlier than planned. There may also be patient concerns regarding randomization to surgical trials, particularly ones assessing nonoperative versus operative management [22]. Investigation of medical devices can be associated with significant learning curves that affect the performance of the device, and there is increased difficulty to blind in these trials. The timing of a trial also must be considered, since devices undergo many modifications during testing in humans, and the device may become outdated while the trial is being completed [21]. While RCTs remain the gold standard, they can be difficult to conduct when investigating surgical interventions or medical devices.

4.5 Phase IV/Device Trials

Once a drug or medical device gains market approval from a regulatory agency such as the Food and Drug Administration (FDA), a phase IV study or post-market study is undertaken to further evaluate the long-term safety. While the FDA has required this since 2007, less than half of these studies are conducted [1]. The goal is to identify less common or long-term adverse events. Phase IV trials may evaluate the drug or device in a cohort with expanded eligibility to simulate real-world situations or investigate cost effectiveness [23]. Although not routinely conducted, this type of study is important, particularly when considering that 20% of drugs will obtain black box warnings, and 4% will be withdrawn due to safety issues during this phase [1].

An example of a phase IV trial is the single-arm prospective trial of the safety and effectiveness of the Magseed® (Endomagnetics, Inc.). The Magseed is a metallic magnetic device for localization of non-palpable breast lesions and received FDA approval in 2016. A post-marketing study of this device was conducted in 107 patients to evaluate for any adverse events and the rate of surgical retrieval of the Magseed along with the lesion of interest. No adverse events occurred, and 100% of the Magseeds were surgically retrieved providing additional confirmatory evidence to the use of this device for localization of non-palpable breast lesions [24].

4.6 Conclusion

Clinical trials are essential for evaluating experimental therapies in human subjects. Although surgical and device trials may have different challenges than traditional drug trials, an understanding of clinical trial design will allow the investigators to be successful at each phase.

References

1. Umscheid CA, Margolis DJ, Grossman CE. Key concepts of clinical trials: a narrative review. Postgrad Med. 2011;123(5):194–204.
2. Le Tourneau C, Lee JJ, Siu LL. Dose escalation methods in phase I cancer clinical trials. J Natl Cancer Inst. 2009;101(10):708–20.
3. Bekele BN, Shen YA. Bayesian approach to jointly modeling toxicity and biomarker expression in a phase I/II dose-finding trial. Biometrics. 2005;61(2):343–54.
4. Wong KM, Capasso A, Eckhardt SG. The changing landscape of phase I trials in oncology. Nat Rev Clin Oncol. 2016;13(2):106–17.
5. Ananthakrishnan R, Menon S. Design of oncology clinical trials: a review. Crit Rev Oncol Hematol. 2013;88(1):144–53.
6. Yuan Y, Hess KR, Hilsenbeck SG, Gilbert MR. Bayesian optimal interval design: a simple and well-performing design for phase I oncology trials. Clin Cancer Res. 2016;22(17):4291–301.
7. Fong PC, Boss DS, Yap TA, et al. Inhibition of poly(ADP-ribose) polymerase in tumors from BRCA mutation carriers. N Engl J Med. 2009;361(2):123–34.
8. Tutt A, Robson M, Garber JE, et al. Oral poly(ADP-ribose) polymerase inhibitor olaparib in patients with BRCA1 or BRCA2 mutations and advanced breast cancer: a proof-of-concept trial. Lancet. 2010;376(9737):235–44.
9. Audeh MW, Carmichael J, Penson RT, et al. Oral poly(ADP-ribose) polymerase inhibitor olaparib in patients with BRCA1 or BRCA2 mutations and recurrent ovarian cancer: a proof-of-concept trial. Lancet. 2010;376(9737):245–51.
10. Brown SR, Gregory WM, Twelves CJ, et al. Designing phase II trials in cancer: a systematic review and guidance. Br J Cancer. 2011;105(2):194–9.
11. El-Maraghi RH, Eisenhauer EA. Review of phase II trial designs used in studies of molecular targeted agents: outcomes and predictors of success in phase III. J Clin Oncol. 2008;26(8):1346–54.
12. Buzdar AU, Ibrahim NK, Francis D, et al. Significantly higher pathologic complete remission rate after Neoadjuvant therapy with Trastuzumab, paclitaxel, and Epirubicin chemotherapy: results of a randomized trial in human epidermal growth factor receptor 2–positive operable breast Cancer. J Clin Oncol. 2005;23(16):3676–85.
13. Buzdar AU, Suman VJ, Meric-Bernstam F, et al. Fluorouracil, epirubicin, and cyclophosphamide (FEC-75) followed by paclitaxel plus trastuzumab versus paclitaxel plus trastuzumab followed by FEC-75 plus trastuzumab as neoadjuvant treatment for patients with HER2-positive breast cancer (Z1041): a randomised, controlled, phase 3 trial. Lancet Oncol. 2013;14(13):1317–25.
14. Buzdar AU, Suman VJ, Meric-Bernstam F, et al. Disease-free and overall survival among patients with operable HER2-positive breast cancer treated with sequential vs concurrent chemotherapy: the ACOSOG Z1041 (Alliance) randomized clinical trial. JAMA Oncol. 2019;5(1):45–50.
15. Jaki T, Vasileiou D. Factorial versus multi-arm multi-stage designs for clinical trials with multiple treatments. Stat Med. 2017;36(4):563–80.

16. Inoue LY, Thall PF, Berry DA. Seamlessly expanding a randomized phase II trial to phase III. Biometrics. 2002;58(4):823–31.
17. Fisher B, Montague E, Redmond C, et al. Comparison of radical mastectomy with alternative treatments for primary breast cancer. A first report of results from a prospective randomized clinical trial. Cancer. 1977;39(6 Suppl):2827–39.
18. Wickerham DL, Costantino JP, Mamounas EP, Julian TB. The landmark surgical trials of the National Surgical Adjuvant Breast and Bowel Project. World J Surg. 2006;30(7):1138–46.
19. Giuliano AE, Hunt KK, Ballman KV, et al. Axillary dissection vs no axillary dissection in women with invasive breast cancer and sentinel node metastasis: a randomized clinical trial. JAMA. 2011;305(6):569–75.
20. Giuliano AE, McCall L, Beitsch P, et al. Locoregional recurrence after sentinel lymph node dissection with or without axillary dissection in patients with sentinel lymph node metastases: the American College of Surgeons oncology group Z0011 randomized trial. Ann Surg. 2010;252(3):426–32.
21. Neugebauer EAM, Rath A, Antoine SL, et al. Specific barriers to the conduct of randomised clinical trials on medical devices. Trials. 2017;18(1):427.
22. Mouw TJ, Hong SW, Sarwar S, et al. Discontinuation of surgical versus nonsurgical clinical trials: an analysis of 88,498 trials. J Surg Res. 2018;227:151–7.
23. Zhang X, Zhang Y, Ye X, Guo X, Zhang T, He J. Overview of phase IV clinical trials for postmarket drug safety surveillance: a status report from the ClinicalTrials.gov registry. BMJ Open. 2016;6(11):e010643.
24. Singh P, Scoggins M, Sahin AA, et al. Effectiveness and safety of Magseed-localization for excision of breast lesions: a prospective trial. Poster presented at: SSO 2019 annual cancer symposium, March 29, 2019; San Diego, CA.

Chapter 5
Defining the Study Cohort: Inclusion and Exclusion Criteria

Emily Z. Keung, Lisa M. McElroy, Daniela P. Ladner, and Elizabeth G. Grubbs

5.1 Introduction

This chapter addresses and reviews the importance of defining the study cohort by means of appropriate eligibility criteria (inclusion and exclusion) in the design of a clinical trial. 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 [1]. 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 target population of the study intervention; and (4) establish the ethical foundation of the study.

E. Z. Keung · E. G. Grubbs (✉)
Department of Surgical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA
e-mail: eggrubbs@mdanderson.org

L. M. McElroy
Section of Transplant Surgery, Department of Surgery, University of Michigan, Ann Arbor, MI, USA

D. P. Ladner
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

T. M. Pawlik, J. A. Sosa (eds.), *Clinical Trials*, Success in Academic Surgery,
https://doi.org/10.1007/978-3-030-35488-6_5

5.2 Overview of Inclusion and Exclusion Criteria

Eligibility criteria consist of a set of inclusion criteria that defines the target patient population of the study intervention and a set of exclusion criteria that refines the study population to remove expected sources of bias and variability [2]. Overly restrictive inclusion and exclusion criteria may limit patient accrual and access to trials and result in studies that fail to capture the heterogeneity of the patient population that will use the intervention after approval. However, heterogeneity of the study population can decrease the accuracy, reliability, and generalizability of the findings of the study [3].

5.2.1 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 carefully selected to minimize the impact of confounding characteristics or variables on the clinical trial outcome(s). 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.

5.2.2 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 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

Grade	ECOG Performance Status
0	Fully active, able to carry on all predisease performance without restriction.
1	Restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature, eg light house work, office work.
2	Ambulatory and capable of all selfcare but unable to carry out any work activities. Up and about more than 50% of waking hours.
3	Capable of only limited selfcare, confined to bed or chair more than 50% of waking hours.
4	Completely disabled. Cannot carry on any selfcare. Totally confined to bed or chair.
5	Dead

Fig. 5.1 ECOG performance status

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 [4]. 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. Patients with poor performance status (PS), defined as a score of 2 on the ECOG rating scale (Fig. 5.1), are often excluded from clinical trials as they tend to have poorer responses to treatment, shorter survival, and possibly greater risk for toxicity than patients with PS scores of 0–1 [5]. 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 on the study objective [6].

5.2.3 Prior Therapy

Prior therapy is often listed in the inclusion/exclusion criteria. Inclusion/exclusion criteria may require study participants to be treatment naïve for multiple reasons. Patients previously treated with standard or investigational therapy may be more likely to have poorer health status, toxicity from treatment, resistance to study drug acquired during prior therapies, or altered response to study treatment and thus may compromise internal validity and need to be excluded from the study cohort [5].

Alternatively, prior therapy may be an inclusion criterion in trial design. Regulatory approval for a study drug may be easier to attain in the second-line setting, or when the established therapy has failed for a given patient [5]. Additionally, patients usually should have the best known available treatment for their disease prior to enrolling in a clinical trial investigating an unproven therapy.

5.2.4 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 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 increase external validity and may facilitate the recruitment of study subjects, they 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 [7, 8].

5.3 Other Considerations in Developing Eligibility Criteria and Participant Accrual for Clinical Trials

5.3.1 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 sample size of an RCT should be determined a priori and reported transparently with scientific justification. Sample size calculations should be performed during the planning phase of a clinical trial and involve the consideration of multiple factors, including the study's objective, type of primary end point to be analyzed, planned analyses, treatment allocation ratio if more patients are to be randomized to one group versus another, allowance of cross-over between study groups, anticipated recruitment rate, estimated number of dropouts, expected underlying event rate in the control group, expected magnitude of treatment effect that the trial is required to detect, the degree of certainty that such detection should occur (the statistical power [1-(the risk of a type II error [beta])]), and the significance level that qualifies as "detected" (the risk of a type I error [alpha]) [9–12]. Clinical trials that aim to detect a small 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, dropout, or non-adherence of study subjects to the intervention [13, 14]. 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 [15]. Using continuous rather than categorical variables to measure outcomes can reduce the sample size required. Other effective approaches to reduce the required sample size 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.

5.3.2 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 will be 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 [8].

5.3.3 Study Cohort Recruitment

There is an array of recruitment methods that are employed for optimal recruitment of subjects into human research studies [8, 16, 17]. For large clinical studies with broad selection criteria, common recruitment methods include the use of advertisements, such as newspaper, radio, television, and internet advertisements.

Additionally, 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 investigator's 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.

5.4 Ethical Considerations

Eligibility criteria for clinical trials are designed to not only define the study population and permit collection of safety and efficacy data specific to the intended population but importantly also protect patients from undue harm [18]. 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 a result of study participation. In addition, all subjects must have the capacity to understand the nature of the study in order to provide informed consent. Standards in place to guide researchers on the ethical inclusion of subjects in research, including special populations, are discussed below.

5.4.1 The Belmont Report

The Belmont Report, issued by the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research in the 1970s, 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: (1) respect for persons (individuals should be treated as autonomous agents, and persons with diminished autonomy are entitled to protection), (2) 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 (3) 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 [19].

5.4.2 Women and Minorities

Federal law requires that women and minorities be included in all clinical research studies, as appropriate for the scientific goals of the work proposed [20]. The NIH Policy and Guidelines on the Inclusion of Women and Minorities as Subjects in Clinical Research [20, 21] 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. In some situations, it may be acceptable to exclude study subjects based on gender or race. Acceptable reasons for exclusion 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. 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.

5.4.3 Inclusion of Children

The Twenty-First Century Cures Act, enacted December 13, 2016, requires the NIH to address the consideration of age as an inclusion variable in research involving human subjects, to identify criteria for justification for any age-related exclusions in NIH research, and to provide data on the age of participants in clinical research studies. The NIH Policy and Guidelines on the Inclusion of Individuals Across the Lifespan as Participants in Research Involving Human Subjects [22] state that individuals of all ages, including children (i.e., individuals under the age of 18) and older adults, must be included in all human subjects research conducted or supported by the NIH, unless there are scientific or ethical reasons not to include them. Acceptable reasons for exclusion include the following: (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 [21]. 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.

5.5 Balancing Patient Protection and Participation

Overly restrictive inclusion and exclusion criteria may limit study accrual, patient access to trials, and generalizability of the study intervention [2]. As regulatory approval is based on data from the enrolled study population, unnecessarily stringent eligibility criteria result in studies that fail to reflect the heterogeneity of the patient population who might benefit from the study intervention [18]. Thus, eligibility criteria must balance ability to discern differences in efficacy and safety outcomes among a less homogeneous group of patient participants and patient safety with applicability and generalizability of study results. In oncology trials, for instance, specific populations are often excluded, such as patients at the extremes of age or those with human immunodeficiency virus (HIV), brain metastasis, history of prior cancer, or organ-system dysfunction [18]. However, there have been recent initiatives to re-examine and modernize clinical trial eligibility criteria for oncology clinical trials to promote greater patient access [23, 24]. The American Society of Clinical Oncology-Friends of Cancer Research established several working groups (Organ Dysfunction, Prior or Concurrent Malignancy, and Comorbidities Working Group, Minimum Age Working Group, HIV Working Group, Brain Metastases Working Group) composed of experts in trial disease and conduct to examine how eligibility criteria could be more inclusive; their recommendations were recently published [25–28].

5.6 Reporting Selection Criteria

To properly interpret and comprehend results of an RCT, readers must understand the study design, conduct, and analysis [29]. Investigators should transparently convey which study subjects were studied and how they were selected. Well-defined, consistent selection criteria allow for ease in reporting study results and allow the reader to understand which patient population the clinical trial intervention applies to and, thereby, which of their patients might benefit most from the studied intervention. Unfortunately, many clinical trials inadequately describe the study population or poorly justify their inclusion/exclusion criteria, making their interpretation difficult and significantly reducing their value [4, 29].

In the 1990s, initiatives to improve the quality of randomized controlled trial reporting led to the development and publication of the CONSORT (Consolidated Standards of Reporting Trials) statement by an international group of clinical trial-

ists, statisticians, epidemiologists, and biomedical editors [30]. Since revised, the CONSORT statement [31] is an evidence-based set of recommendations for the standardized reporting of results from RCTs in a complete and transparent fashion that includes a 22-item checklist (Fig. 5.2) and flow diagram (Fig. 5.3) and which assists the critical appraisal and interpretation of study results.

Section/topic	Item no	Checklist item	Reported on page no
Title and abstract			
	1a	Identification as a randomised 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 were actually 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 assessing outcomes) and how	
	11b	If relevant, description of the similarity of interventions	
Statistical methods	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. 5.2 CONSORT 2010 checklist of information to include when reporting a randomized trial. We strongly recommend reading this statement in conjunction with the CONSORT 2010 Explanation and Elaboration for important clarifications on all the items. If relevant, we also recommend reading CONSORT extensions for cluster randomized trials, non-inferiority and equivalence trials, non-pharmacological treatments, herbal interventions, and pragmatic trials. Additional extensions are forthcoming: for those and for up to date references relevant to this checklist, see www.consort-statement.org.

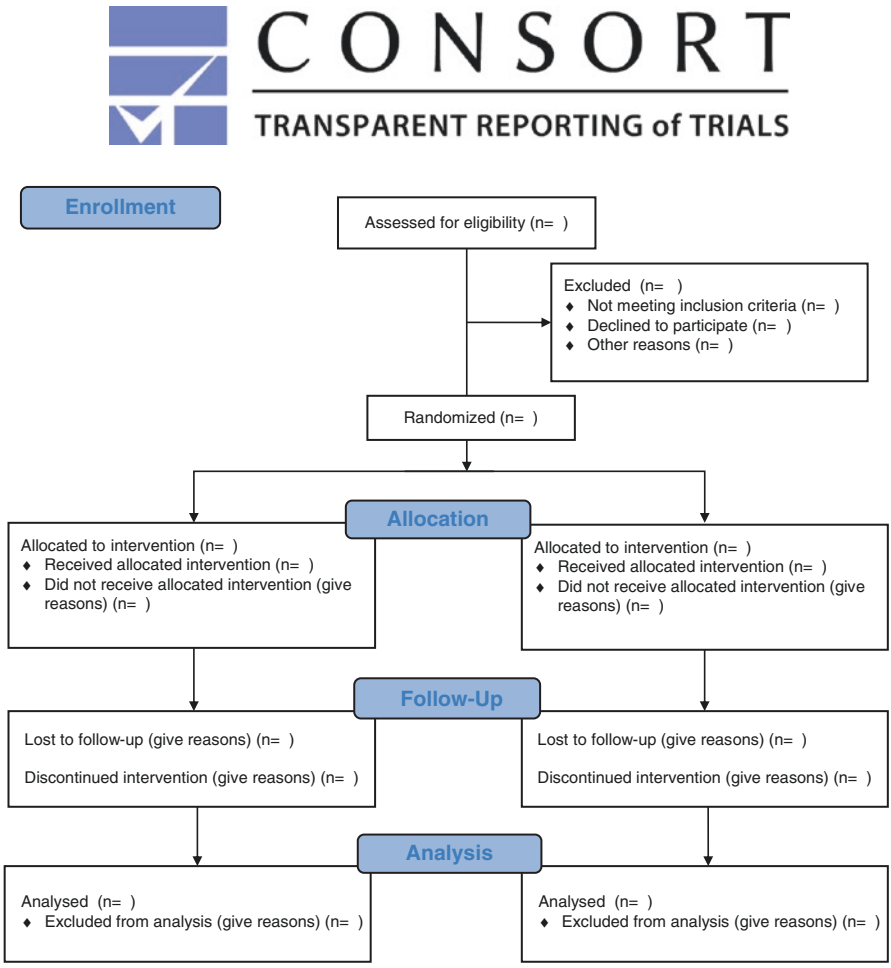


Fig. 5.3 CONSORT 2010 flow diagram

5.7 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 [6].

5.8 Conclusion

This chapter focused on the importance of defining the study cohort of a clinical trial by means of carefully considered eligibility criteria. Eligibility criteria should be developed and defined during the planning phase of a clinical trial with the study goals in mind. Inclusion and exclusion criteria determine who is eligible to participate in a clinical trial. Eligibility criteria should represent the intervention's target population while protecting the safety of trial participants and be sufficiently narrow to ensure internal study validity. However, overly restrictive eligibility criteria can hinder trial accrual, restrict understanding of the intervention's risk-benefit profile, and limit the applicability and generalizability of study results.

References

1. ClinicalTrials.gov. Learn about clinical studies. https://clinicaltrials.gov/ct2/about-studies/learn. Accessed 25 July 2019.
2. Underwood RT. Basic design considerations. In: Chow S-C, Liu J-P, editors. Design and analysis of clinical trials: concepts and methodologies. 3rd ed. New York: John Wiley & Sons; 2014. p. 85–115.
3. Kim ES, Bernstein D, Hilsenbeck SG, et al. Modernizing eligibility criteria for molecularly driven trials. J Clin Oncol. 2015;33(25):2815–20.
4. Van Spall HGC, 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(11):1233–40.
5. Brody T. Inclusion/exclusion criteria, stratification, and subgroups—part I. In: Brody T, editor. Clinical trials. 2nd ed. Amsterdam: Elsevier; 2016. p. 83–113.
6. Chin RY, Lee BY. Patient selection and sampling: key components of clinical trials and programs. In: Principles and practice of clinical trial medicine. Boston: Academic; 2008.
7. Nottage M, Siu LL. Principles of clinical trial design. J Clin Oncol. 2002;20(18 Suppl):42S–6S.
8. Szklo M. Population-based cohort studies. Epidemiol Rev. 1998;20(1):81–90.
9. Ellimoottil C, Vijan S, Flanigan RC. A primer on clinical trial design. Urol Oncol Semin Orig Investig. 2015;33(3):116–21.
10. Stock E, Biswas K. Sample size calculation. In: Itani KMF, Reda D, editors. Clinical trials design in operative and non operative invasive procedures. New York: Springer International Publishing; 2017. p. 141–50.
11. Pocock SJ, Clayton TC, Stone GW. Design of major randomized trials: part 3 of a 4-part series on statistics for clinical trials. J Am Coll Cardiol. 2015;66(24):2757–66.
12. Browner W, Newman T, Cummings S, Hulley S. Estimating sample size and power. In: Hulley S, Cummings S, Browner W, Grady D, Hearst N, Newman T, editors. Designing clinical research: an epidemiologic approach. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2001. p. 65–84.
13. Schulz KF, Grimes DA. Sample size calculations in randomised trials: mandatory and mystical. Lancet. 2005;365(9467):1348–53.
14. Kirby A, Gebski V, Keech AC. Determining the sample size in a clinical trial. Med J Aust. 2002;177(5):256–7.

15. Caldwell PHY, Hamilton S, Tan A, Craig JC. Strategies for increasing recruitment to randomised controlled trials: systematic review. PLoS Med. 2010;7(11):e1000368.
16. Treweek S, Pitkethly M, Cook J, et al. Strategies to improve recruitment to randomised trials. Cochrane Database Syst Rev. 2018;2(2):MR000013.
17. Itani KMF, Reda DJ. Clinical trials design in operative and nonoperative invasive procedures. Clin Trials Des Oper Non Oper Invasive Proced. 2017;112:1–495.
18. Beaver JA, Ison G, Pazdur R. Reevaluating eligibility criteria—balancing patient protection and participation in oncology trials. N Engl J Med. 2017;376(16):1504–5.
19. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. The Belmont Report. http://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/#. Published 1979. Accessed 25 July 2019.
20. U.S. Department of Health and Human Services. National Institutes of Health. NIH policy and guidelines on the inclusion of women and minorities as participants in research involving human subjects. https://grants.nih.gov/policy/inclusion/women-and-minorities.htm. Accessed 25 July 2019.
21. National Institutes of Health. Guidelines for the review of inclusion on the basis of sex/gender, race, ethnicity, and age in clinical research. https://grants.nih.gov/grants/peer/guidelines_general/Review_Human_subjects_Inclusion.pdf. Accessed 25 July 2019.
22. U.S. Department of Health and Human Services. National Institutes of Health. NIH policy and guidelines on the inclusion of individuals across the lifespan as participants in research involving human subjects. https://grants.nih.gov/policy/inclusion/lifespan.htm. Accessed 25 July 2019.
23. Jin S, Pazdur R, Sridhara R. Re-evaluating eligibility criteria for oncology clinical trials: analysis of investigational newdrug applications in 2015. J Clin Oncol. 2017;35(33):3745–52.
24. Penel N, Lebellec L, Vanseymortier M. Reappraisal of eligibility criteria in cancer clinical trials. Curr Opin Oncol. 2018;30(5):352–7.
25. Lichtman SM, Harvey RD, Damiette Smit M-A, et al. Modernizing clinical trial eligibility criteria: recommendations of the American Society of Clinical Oncology-Friends of Cancer research organ dysfunction, prior or concurrent malignancy, and Comorbidities Working Group. J Clin Oncol. 2017;35(33):3753–9.
26. Gore L, Ivy SP, Balis FM, et al. Modernizing clinical trial eligibility: recommendations of the American Society of Clinical Oncology-friends of Cancer research minimum age working group. J Clin Oncol. 2017;35(33):3781–7.
27. Uldrick TS, Ison G, Rudek MA, et al. Modernizing clinical trial eligibility criteria: recommendations of the American Society of Clinical Oncology-friends of Cancer research HIV working group. J Clin Oncol. 2017;35(33):3774–80.
28. Lin NU, Prowell T, Tan AR, et al. Modernizing clinical trial eligibility criteria: recommendations of the American Society of Clinical Oncology-friends of Cancer research brain metastases working group. J Clin Oncol. 2017;35(33):3760–73.
29. Moher D, Schulz KF, Altman D. CONSORT group (consolidated standards of reporting trials). The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. JAMA. 2001;285(15):1987–91.
30. Begg C, Cho M, Eastwood S, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. JAMA. 1996;276(8):637–9.
31. The CONSORT statement. http://www.consort-statement.org/. Accessed 25 July 2019.

Chapter 6
Pragmatic Trials and Approaches to Transforming Care

Peter G. Stock, Rita Mukhtar, Hila Ghersin, Allison Stover Fiscalini, and Laura Esserman

When considering medical interventions and tools for learning and advancing a field, one perspective is to test a drug, device, or intervention in a highly controlled setting, where inclusion and exclusion criteria are very strict. Such criteria can limit the impact of bias and confounders, while providing rigor in assessing the impact of an intervention. However, when evaluating data from a clinical trial and assessing whether these criteria apply to the patient who sits in front of you, this approach creates a challenge: such strict criteria often mean that the person for whom you want to apply "the evidence" is not appropriate. For this reason, there has been a move to conduct more pragmatic trials that are designed to test the effectiveness of the intervention in broad routine clinical practice. Often, interventions that show a dramatic impact in the setting of a clinical trial fail to be effective in broader settings. This phenomenon is called regression to the mean [1]. So one way to try to approach the assessment of drug, device, and surgical interventions is to evaluate them using a pragmatic trial approach.This establishes a broader base for the intervention and, prospectively, you can identify the subgroups where the interventions could be found to be more effective. This improves applicability of the results.

Schwartz and Lellouch describe a pragmatic trial as one which informs a clinical or policy decision by providing evidence for adoption of the intervention by real-world clinical practice [2]. In a comprehensive overview, Ford and Norrie [3] describe multiple features of a pragmatic trial, including (1) a design to show the real-world effectiveness of the intervention in broad patient groups, (2) approaches to improve the effectiveness of the intervention, (3) inclusion of a population that is relevant for the intervention and a control group treated with an acceptable standard of care, and (4) outcomes which are meaningful and conducted and analyzed at a high standard of quality.

P. G. Stock (✉) · R. Mukhtar · H. Ghersin · A. Stover Fiscalini · L. Esserman
Department of Surgery, University of California, San Francisco, CA, USA
e-mail: Peter.Stock@ucsf.edu; laura.esserman@ucsf.edu

T. M. Pawlik, J. A. Sosa (eds.), *Clinical Trials*, Success in Academic Surgery,
https://doi.org/10.1007/978-3-030-35488-6_6

In this chapter, we describe four different examples of how pragmatic approaches are being used to generate evidence, evolve trial designs, and change practice. The first is a unique trial that led to the acceptance of solid organ transplantation in people infected with HIV. The second is a PCORI (Patient Centered Outcomes Research Institute)-funded trial, which encourages the inclusion of a broad population to generate evidence. The third is an approach to modernize guidelines for inclusion or selection for organ transplantation. And finally, we discuss how changes in practice require a change in how we approach trial design. When clinicians begin to act on data and early outcome measures, traditional randomization and follow-up is impossible, thus requiring creative approaches to continue to advance the field. For this example, we will briefly describe trials in the neoadjuvant breast cancer setting and the challenges posed as early endpoints prove to be highly prognostic.

6.1 HIV and Transplant: Safety and Efficacy of Solid Organ Transplantation in the HIV-Infected Recipient

A pragmatic approach was necessary in the design of an NIH multicenter trial to determine the safety and efficacy of liver and kidney transplantation in people infected with HIV. The impetus for the trial was to provide evidence for whether or not to allow transplantation in an HIV-infected person and whether or not it should become a clinical care standard. At the time the trial was initiated, HIV positivity was a strict contraindication to transplantation in the vast majority of transplant centers. The need to establish effectiveness of this approach was essential. Organ donors are a scarce resource, and the number of people who could benefit from transplantation far exceeds the availability of deceased donor organs. Using an organ donation in a setting where it might have a high risk of failure was therefore considered unethical. However, as HIV became a chronic disease and organ failure emerged as the major life-threatening illness in HIV-infected patients, an ethical dilemma emerged. Should patients classically excluded from trials of and registries for transplantation continue to be excluded? For that reason, a trial was designed which explicitly tested the safety of transplantation in a population typically excluded from trials and care. The NIH trial met all the features of pragmatic trials as outlined by Ford and Norrie [3]. The particular challenge the investigators faced with unique funding issues and recruitment of patients and investigators is also described below. Of interest, funding of standard of care as part of trials is a common challenge faced by many pragmatic trials.

Toward the late 1990s, it became apparent that people infected with the human immunodeficiency virus (HIV) were no longer dying from progression of HIV to AIDS. HIV infection was effectively controlled with combined antiretroviral therapy (cART) and so had evolved into a chronic condition. However, based on the

comorbidities associated with HIV, there was an increasing incidence of end-stage kidney and liver disease in people with well-controlled HIV. At the time, due to concerns regarding potential exacerbation of an already immunocompromised state with the immunosuppression required for liver or kidney transplantation, people infected with HIV were not considered candidates for transplantation. Concurrently, our transplant center at the University of California, San Francisco (UCSF) was seeing an increasing number of HIV patients referred for liver and kidney transplantation. Notably, HIV nephropathy had become the third most common cause of end-stage renal failure in young people of African descent. Hepatitis and HIV shared many of the same risk factors and end-stage liver disease had become a significant cause of death in people infected with HIV. It is in that setting that we faced the challenge of developing a clinical trial to assess the safety and efficacy of solid organ transplantation in HIV-infected individuals.

6.1.1 Funding the Trial

Obtaining funding for a large multicenter trial studying the safety and efficacy of solid organ transplantation was particularly challenging, in that third-party payers did not provide reimbursement for the clinical costs of transplants in people with HIV infection. In addition to our center, several transplant programs across the state and country were recognizing the increasing need for solid organ transplantation in this population, and community activists facilitated a meeting of third-party payers in San Francisco. During that public meeting, our transplant team presented data suggesting that it was time to move forward with a clinical trial in light of the increasing number of well-treated HIV-infected people with end-stage liver and kidney disease. We also discussed the importance of having a standardized approach and protocol, so that the results regarding safety and efficacy of transplantation in the HIV-infected recipient could be generalized to transplant centers across the country. Interestingly, activists that were present at this public hearing were concerned that having a formal study protocol translated to a "no payment" verdict from third-party payers, who do not reimburse "experimental" procedures. This skepticism by community activists had its genesis from the poor initial response to the HIV epidemic by the National Institutes of Health and lack of initial funds for research in the early days of the HIV crisis. The local community activists staged a protest in front of the UCSF hospital, despite the fact that we were advocating for transplantation in HIV-infected people with end-stage liver and kidney disease. However, their concern shed light on the complexity of the issues and the need to support a pilot. Suffice it to say that the activists were instrumental in terms of procuring funding from the state of California for a pilot safety and efficacy trial; the data from the pilot trial formed the preliminary data necessary to secure funding for the large NIH multicenter trial (HIV-TR). The NIH multicenter HIV-TR was a prospective non-randomized, unblended safety and

efficacy trial which followed 175 kidney transplant and 125 liver transplant recipients and followed the patients from active listing on the UNOS waiting list through transplantation.

6.1.2 Recruitment of Study Participants

This prospective study was designed to include HIV-infected patients with end-stage liver or kidney disease with well-controlled HIV; the goal was to establish safety and efficacy so HIV would no longer be viewed as a contraindication to transplantation. Inclusion/exclusion criteria therefore required CD4+ T-cell counts >200 cells/ml for kidney transplant recipients and >100 cells/ml for liver transplant recipients (lower counts in chronic liver disease anticipated as a result of splenic sequestration), levels that were viewed as the threshold of T-cell counts for fighting opportunistic infections. Kidney transplant recipients had to have undetectable HIV-1 RNA, whereas liver recipients could have detectable HIV-1 RNA if a fully suppressive cART regimen could be maintained following liver transplant. The less rigid requirement for liver transplant recipients related to the fact that some components of cART were hepatotoxic and therefore had to be stopped to prevent exacerbation of the chronic liver disease. Patients with a history of opportunistic infections or cancers without effective therapeutic options were also excluded from the trial. These fairly broad criteria were used as they reflected good control of HIV-1 infection on cART and a better likelihood of being able to tolerate the additional immunosuppression required following transplantation. Furthermore, based on our preliminary data from a pilot trial using the same inclusion/exclusion criteria, liver and kidney transplant recipients tolerated the procedure and the required immunosuppression without loss of HIV control and progression of HIV to AIDS. As the goal of the trial was to encourage transplantation into real-world practice, precise immunosuppressive protocols were not dictated—the more restrictive the protocol, the less likely we would be able to enroll patients and have the support of transplant centers across the country. Furthermore, we wanted to be able to evaluate the relative efficacy of the various strategies as a means of improving future outcomes.

6.1.3 Recruitment of Investigators

Although there was no shortage of HIV-infected patients who were in need of liver or kidney transplantation and who met the broad entry criteria, a bigger challenge was identifying centers in each geographic region to participate in this trial. A major obstacle was appropriate safety concerns related to risks of HIV transmission to the

teams performing the transplants. Needle sticks are not uncommon in the operating room and, of course, can happen with routine blood draws as part of routine management. With the help of the HIV providers, antiretroviral regimens which are effective against the strain of HIV for each recipient were made available for immediate use in the event of a needle stick. Furthermore, participation in the surgical procedure was optional for residents, technicians, and nurses—although at UCSF we did not have a single instance where the staff did not want to participate. There were a few needle sticks in the operating room, and antiretroviral preparations were immediately available. Fortunately, to our knowledge, there were no transmissions of HIV to healthcare workers in this study in long-term follow-up. A second obstacle to participation was center concern for poorer outcomes in this higher-risk group of transplant recipients. Transplantation is one of the most regulated fields and center-specific results are monitored and made available in the public record. If center results drop below a given threshold, transplant centers are at risk for losing referrals, as well as insurance coverage. For this reason, we enrolled centers that could take the additional risk without impacting center-specific results in a significant way. By having a large number of centers with a geographic spread, we were able to accomplish this goal. Finally, in order to encourage center participation, we welcomed participation of basic scientists in each center to become involved with the mechanistic studies examining the impact of immunosuppression on the immune response in HIV-infected individuals. The ability to engage the basic scientists in the study enabled broader support for the study and helped to encourage participation in the study. As well, it improved our application for funding because we included state-of-the-art mechanistic studies. By including the entire transplant team, from basic scientists to pharmacologists to clinicians, we increased our ability to get funding, and we improved our ability to recruit centers of excellence across the country and therefore enhanced broad participation in this trial.

6.1.4 The Trial, Outcomes, Impact on Practice, and Future Directions

The strategy to recruit 20 centers across the country in areas with a broad geographic distribution in regions with a high incidence of HIV infection was successful. The trial enrolled 175 kidney transplant recipients and 125 liver transplant recipients who met the inclusion/exclusion criteria. This cohort of HIV-positive transplant recipients was compared to matched HIV-negative registry controls. For the kidney transplant recipients, the cohort of HIV-positive recipient was not only compared to matched registry controls but also a subset of HIV-negative kidney transplant recipients over the age of 65. This later subset was a chosen since this cohort, like HIV-positive kidney transplant recipients, is a selected group with higher risk factors but considered candidates for transplantation. We felt this was an import comparator, in that deceased donor

organs are a scarce resource, and we were hopeful the trial would demonstrate that the utilization of this scarce and valuable resource in people with HIV infection was comparable to the HIV-negative recipients. Similarly, the liver transplant cohort was compared to matched SRTR controls. Without going into the details, transplantation was associated with survival benefit for HIV-infected liver recipients with model for end-stage liver disease score (MELD) greater than or equal to 15 ($p < 0.0002$). In HIV-positive kidney recipients, unmatched and risk-matched analyses indicated a marginally significant hazard ratio (HR) for graft loss (unmatched 1.3 ($p = 0.07$) and risk-matched 1.5 ($p = 0.052$)); no significant increase in risk of death was observed [4]. More details on the four publications from this trial will not be described in this chapter, but should be reviewed for details [5–7]. Critically, however, the results of this pragmatic trial have resulted in the removal of HIV positivity as a contraindication to transplantation.

Although the safety and efficacy of solid organ transplantation in the HIV-infected recipient was the primary goal of HIV-TR, the secondary analyses and mechanistic studies continue to impact strategies for providing better care in this cohort. Interestingly, there was a higher than expected incidence of rejection in both liver and kidney recipients, and strategies for decreasing rejection have been developed based on the mechanistic studies. Surprisingly, certain antiretroviral drugs may have an impact on blocking the immune response, and certain immunosuppressive drugs may have an impact on decreasing the HIV viral reservoir. Of equal significance, national laws have been changed to permit the utilization of organs from HIV-positive deceased donors for use in HIV-positive recipients. At the end of the HIV-TR trial, there were over 350 HIV-positive recipients on the national waiting list, and using the HIV-positive organs in these recipients will permit transplantation in a more expeditious manner, further improving results in HIV-infected people in need of solid organ transplantation.

6.2 WISDOM: Women Informed to Screen Depending on Measures of Risk

Created as part of the Affordable Care Act, the Patient-Centered Outcomes Research Institute (PCORI) has a category for funding pragmatic trials that are patient-centered and address issues that are common and important to the population as a whole. PCORI has several stipulations for their pragmatic clinical trials. One is that the trials must be patient-centered and broadly inclusive of the populations affected by the condition being studied. Another stipulation is that the results be generated in 5 years. Patient-centered outcomes must be highlighted in the study, with patient advocates playing a central role on the study team from the beginning. As well, investigators must have a plan to address the gap between the generation of evidence and adoption of results. One of the more challenging barriers in performing this and other pragmatic trials is that funding for clinical interventions must be covered as part of clinical care rather than with PCORI funds.

Breast cancer screening in the USA is extremely contentious, and there are public disagreements among major professional organizations regarding appropriate guidelines. The radiology societies recommend annual screening starting at 40, the US Preventive Services Task Force recommends biennial screening for women 50–74 years of age, and other organizations fall in-between. Women are caught in the middle. But breast cancer is not one disease, and women do not have the same risk factors, so it is extremely unlikely that a one-size-fits-all approach is best. Furthermore, screening itself is not harmless. Among its known harms are high rates of false-positive recall and biopsy and overdiagnosis and overtreatment [8].

The Athena Breast Health Network was launched initially as a collaboration between all the University of California medical centers and Sanford Health with the goal of integrating care and research across the spectrum from screening and prevention to treatment and survivorship. One of the key gaps in care that the network was keen to address was to integrate into the screening process the vast advances in our understanding of the biology of breast cancer and breast cancer risk assessment. The WISDOM study (Women Informed to Screen Depending on Measures of risk) is our response [9]. WISDOM aims to change the paradigm for breast cancer screening, by testing a new personalized approach to screening and prevention against one-size-fits-all annual mammography. By starting with risk assessment, the goal is to personalize a plan for prevention and screening and learn how to do more for those that need it and less for those that do not. Reflecting the pragmatic design, WISDOM has few restrictions on enrollment. We encourage all women with no history of breast cancer between the ages of 40 and 74 to go to wisdomstudy.org and join the study. The aim is to enroll 100,000 women.

One of the more interesting aspects of WISDOM is its "preference-tolerant" design that was developed in partnership with our patient advocates. It is a critical component to ensure a pragmatic approach to enrollment. Women are encouraged to be randomized. We explain that randomization is the best way to learn and answer the question about safety. Most women spend 35 years of their life getting screened for breast cancer. Over the course of 5 years, the results will be available, and all participants will then know the best option. However, if women feel strongly about one option or the other, they can choose to join the observational cohort and choose the arm in which they want to participate. They have the choice of joining either the annual or personalized arm. That way no person is excluded from joining the trial. Rather than including only those who are willing to be randomized, we explicitly designed the study to have a randomized component as well as an observational arm. This strategy means that there is no barrier to enrollment. Persons who feel strongly can choose the arm in which they want to participate rather than be excluded. This addresses a very important bias. Those who are not willing to be randomized can still be represented. It enables learning about patient preferences and, for those who are not at equipoise, which arm of the study they prefer. As well, pragmatic approaches to inclusion and exclusion ensure that adoption is broader. There are many questions that can be answered, whether a person is randomized or chooses an intervention (rates of intervention, e.g., biopsy). There are some questions that are best answered with a randomized cohort (safety and efficacy) and

some where choice is offered (e.g., assessing patient preference). Finding ways to include as broad a population as possible also affords the ability to evaluate each cohort to understand the inherent biases in recruiting only a randomized cohort.

Another fundamental approach to ensuring broad representation among various populations is ensuring ease of access to the trial. WISDOM does not require patients to go to a specific clinical site for study visits. The majority of enrollment and participation occur online, and participants continue to see their regular providers for their care, including for their mammograms. As well, we have been able to partner with a company that can send spit kits for genetic testing directly to a person's home. The trial is now translated into Spanish, and we are expanding across the country and targeting specific populations to increase diversity of subjects enrolled.

6.2.1 Funding the Trial and the Interventions

There is currently a gap in funding of pragmatic comparative effectiveness trials, in that clinical procedures that are experimental are not covered by payers, but similarly are not covered by traditional funding agencies, PCORI included. In the case of WISDOM, the cost of the genetic testing required for the risk-based screening approach fell into this funding nether region. We pursued a Coverage with Evidence Development (CED) policy when designing WISDOM, which was inspired by Medicare's CED policy. CED and CEP (Coverage with Evidence Progression) support evidence development within the healthcare system for new applications of existing technologies. For example, the use of a genetic testing for women with a family history of breast cancer is already standard of care; genetic testing is not experimental. And with changes in the legal framework and advances in the technology of next-generation sequencing, the quality of testing has improved (there are now nine genes where mutations are associated with a significant increase in the risk of breast cancer) as the costs (from some companies) have dramatically dropped, where the test is approximately the cost of a mammogram. So, within the WISDOM study, the test is used for all women (regardless of family history) in the personalized screening arm to identify the highest-risk women and maximize the potential for prevention within the context of a clinical trial.

In the years prior to the study's inception and during the initial recruitment years, the WISDOM team worked to adapt the CEP framework with private payers, such as Blue Shield of California. By working with private payers, whose members would fall into the study population, the WISDOM study could make the case for evidence development in a real-world setting befitting of the pragmatic trial. The WISDOM CEP framework expanded too after endorsement from the Blue Cross Blue Shield Association [10]. Nearly a dozen regional Blue Cross Blue Shield plans agreed to participate in WISDOM as a pilot for CEP within their organization. The intent was to offer the trial broadly, but challenges to implementation were considerable.

The intent of CEP-based clinical research is to deliver study services nationally by developing scalable billing and study procedures that reduce administrative burden. The WISDOM study uses a unique national provider identification (NPI) and tax identification number (TIN) for all billing services that can be used as the study scales [11]. This has markedly reduced administrative overhead and has centralized claim submission as the study has expanded outside California. Although the WISDOM study sites include numerous medical centers, all study-related clinical services are billed by a single NPI under the University of California Office of the President's Center for Health Quality and Innovation (CHQI). This minimized system and workflow changes to the payer systems, allowing for more rapid adoption with minimal resources.

To ensure a maximal number of study participants could participate in the study, the WISDOM study approached employers with self-insured plans and Medicaid as well. In general, health plans are not generally staffed or configured to support direct "marketing" of a benefit like WISDOM to their large self-funded clients. The WISDOM team directly engaged with larger self-insured and flex-funded employers. This active collaboration allowed employers to approach their TPA to implement CEP for their employees and dependents. These benefits would be administered by the third party administrator (TPA) on behalf of and with the support of the employers. By working with these groups, which often can dictate their own plan coverage, the WISDOM study was able to expand the eligible study population through CEP.

6.2.1.1 Recruitment of Study Participants

Another aspect of pragmatic trials is to address the implementation and the framework used for implementation. In order to broadly recruit for the study, we needed to make the trial available to everyone in the population. In order to get companies and payers to agree to partner with us, we could not restrict accrual to specific sites but needed to broaden our ability to recruit patients. This also required us to be practical in how we recruited patients and collected data.

We used modern tools to set up the study, using a cloud-based platform based on Salesforce (Salesforce Inc., San Francisco) and their integrated analytics for real-time capture of trial accrual. We also partnered with a number of companies that are modernizing the approach to delivering tests (e.g., Color Genomics) and collecting imaging data (Life Image).

6.2.1.2 Data Collection Is Pragmatic

Enrollment and participation is online. Participants provide electronic consent and complete online questionnaires in their study portal. Saliva-based genetic testing kits are mailed to the participant's home and mailed back directly to the lab at Color. All study documents and communications are completed online. Mammography

records are obtained through a variety of channels, including direct integration with local electronic medical records, Mammosphere patient portal, and electronic fax. Mammosphere has partnered with WISDOM to provide an electronic portal for digital health record ascertainment and exchange between providers, in response to recent HIPAA requirements to provide electronic health records to requesting patients. The pragmatic approach to data collection and participation in WISDOM has enabled broad recruitment and access to the study. Many traditional barriers to trial participation, including transportation, availability of trials at local care facilities, and inconvenience, are all overcome by our pragmatic trial recruitment approach.

6.2.1.3 The Trial, Outcomes, Impact on Practice, and Future Directions

The WISDOM study is ongoing, and we have recruited over 22,000 women at the time of writing this chapter. In order to reduce the time from study results to having an impact on guidelines, we have employed a stakeholder-engagement model and have recruited leaders from every sector, including the major guideline and quality standard bodies, payers (insurers and self-insured companies), advocates, researchers, clinical leaders, and technology partners. All stakeholders are encouraged to advise and engage in evaluating the simulations of results in an effort to reduce the time it will take to accept and adopt finding from the study. This multi-stakeholder model enables participation from organizations like the National Committee for Quality Assurance (NCQA), which independently accredits insurance companies through the HEDIS measure. A higher HEDIS score indicates higher quality, so health plans may be concerned about participating if it lowers their HEDIS score for breast cancer screening. However, the WISDOM protocol is consistent with HEDIS requirements, as the HEDIS measures were recently changed to accommodate the range of guidelines [12]. If the study demonstrates the value of risk-based screening and identifies opportunities to further adapt screening frequencies based on risk, current HEDIS rules may need revision. The multi-stakeholder model enables organizations like the NCQA, US Preventive Services Task Force, and American Cancer Society to incorporate findings from CEP trials into their measure definitions or guidelines.

Stakeholders have also become champions as the study expanded nationwide. For example, the WISDOM study worked with various plan leadership, from organizations like the Blue Cross Blue Shield Association, to align strategic goals. In 2017, the Blue Cross Blue Shield Association announced the CEP framework to endorse expansion of the WISDOM study to additional participating pilot Blue Cross Blue Shield companies [10]. Our broad stakeholder group meets annually in-person to review study progress, to discuss challenges and opportunities, and to review modeled study outcomes in order to advise on study implementation and build a strategy for dissemination.

The WISDOM study incorporates a number of important pragmatic trial elements and, if successful, could serve as a model for other screening trials. It includes both randomization and individual choice. It includes a coverage with evidence model that allows payers to contribute to the advancement of the field and accelerate

change; WISDOM also includes a strategy whereby advances in the field are incorporated as the trial proceeds, ensuring that the trial will not be out of date when the results are published. Results are collected in real time, using a modern software platform. And finally, the stakeholder model helps guideline and policy makers to have a seat at the table and discuss and view results in real time to decrease the gap between release of results and change in practice. WISDOM is therefore an excellent example of how pragmatic trial design choices can influence policy, guidelines, and the speed of adoption of results.

6.3 Transplantation in the Setting of Breast Cancer History: A Pragmatic Approach to Changing Policy Through Modernization of Inclusion and Exclusion Criteria

Pragmatic approaches should also allow us to address important opportunities to improve policies and creative approaches to generate confirmatory data. For patients with unique clinical problems or several concurrent conditions, there is often a lack of prospective data to use for guiding management decisions. When existing data generated in highly controlled settings are applied to patients in these particular situations, care can be suboptimal. One potential area that could benefit from a pragmatic approach is the intersection between organ transplantation and management of underlying disease states such as cancer.

Breast cancer is the most common malignancy diagnosed in women annually, and therefore it is a common comorbid condition in patients with end-stage renal disease (ESRD). Of the two million patients who have ESRD worldwide, about 5% will also have a diagnosis of breast cancer [13]. Historically, pre-transplant malignancies like breast cancer have been viewed as a contraindication to kidney transplantation. Uncertainty about cancer recurrence and subsequent mortality, coupled with concerns regarding the judicious use of scarce donor organs in patients who have a limited prognosis, have prevented this patient population from immediate kidney transplantation [14]. Research from the Israel Penn International Transplant Tumor Registry reported that patients with pre-existing breast cancer have high recurrence rates after transplantation, ranging from 5.4% to 63.6% [15, 16]. Based on the results from studies like this, organizations and institutions have implemented waiting times ranging from a minimum of 2–5 years to ensure that breast cancer recurrence is unlikely to occur prior to receiving kidney transplantation.

However, this mandatory years-long waiting period leads to increased morbidity and potentially mortality. Patient with ESRD who wait for kidney transplantation rely on dialysis, which has a 5-year survival rate of 35% [13]. Many of the current transplantation guidelines recommend equivalent waiting times for patients who are at low and high risk of breast cancer recurrence and fail to account for tumor characteristics like biomarkers and grade—factors that inform this risk. While the field of breast cancer has changed dramatically and our ability to predict outcome and predict response to therapy has greatly improved, the integration of this information into other fields has lagged. The

example of how the advances in science of management of breast cancer should influence transplant decisions (for donors and recipients) is an example of an important but unique clinical situation for which a pragmatic approach would speed up the implementation of diagnostics that could help patients with ESRD avoid potentially unnecessary waiting times, substantial morbidity, and potentially increased mortality.

6.3.1 *Applying Scientific Advances to the Classification of Breast Cancer to Inform Transplant Eligibility*

Recently, a small case series showed how wait time was eliminated for two patients with ESRD and prior breast cancer [17]. These patients were able to receive immediate kidney transplants based on recurrence risk scores generated by genomic expression profiling assays (e.g., Oncotype DX and MammaPrint). Individualizing patients' risk of breast cancer recurrence by using molecular assays in combination with tumor pathology could potentially impact a large number of patients who might normally be excluded from life-saving interventions based on data that are not personalized. Additionally, it is the most aggressive, high-risk cancers that are likely to be driven by immunosuppression. In triple-negative tumors, for example, immune infiltrates are more common, and reversal of local immunosuppression has dramatically improved that chance of complete response [18]. On the other hand, molecularly low-risk tumors are "cold" and lack immune infiltrates and are not likely to be influenced by immunosuppression. Since the risk and timing of recurrence differ by molecular risk status, tailoring waiting time to breast cancer subtype is a more reasonable approach to avoid unnecessarily delaying transplantation. Additionally, our understanding of prognosis is improved by incorporation of response to therapy. Guidelines should reflect such modern approaches to breast cancer management [19, 20].

6.3.2 *Lack of Standards and Consensus Among Transplant Surgeons and Nephrologists*

Additionally, a survey of 129 transplant surgeons and nephrologists from 14 countries and 32 states confirms that the existing guidelines for managing kidney transplantation decisions in patients with breast cancer are inadequate. 74.8% of respondents felt current guidelines are not sufficient to inform their decision-making if faced with a potential kidney transplant candidate who has a breast cancer diagnosis. Of the providers that were surveyed, 27% didn't think there were standard guidelines for this patient population, and 9% weren't aware of the guideline recommendations. Furthermore, transplant management for patients who have a history of breast cancer varied among providers and differed depending on geographic location and beliefs about immunosuppression's role in breast cancer recurrence. Lack of consensus in the transplant community contributes to variable access to potentially life-saving organs for breast cancer patients.

6.3.3 Recommendations for the Future and Impact on Practice: Not Every Question Requires a Trial

Conducting pragmatic trials for patients with ESRD and breast cancer could inform new guidelines in the field and improve patient outcomes. Alternatively, new criteria could be set and real-world evidence can be used to determine the outcomes. Especially given that transplant patients are in registries, any patients with a history of breast cancer can be followed closely. Results at 5 years would be sufficient to inform the field. This is the ultimate pragmatic study.

6.4 I-SPY TRIALS: Investigation of Serial Studies to Predict Your Therapeutic Response with Imaging and Molecular AnaLySis

A final example of a pragmatic trial design is the I-SPY TRIAL, which uses an adaptive strategy to enable rapid learning in a high-risk setting to facilitate the identification of optimal drug combinations to cure patients. An *adaptive design* is defined as "a **design** that allows modifications to the trial and/or statistical procedures of the trial after its initiation without undermining its validity and integrity" [21]. The purpose is to make clinical trials more flexible, efficient, and fast. Drug development in oncology usually focuses on the metastatic setting where phase I (safety and dose finding), II (signal-generating), and III (efficacy) trials are conducted in people with advanced disease before proceeding to the early treatment setting. The current approach in trials is to test one drug at a time that patients are still curable (high risk for recurrence but still early stage and curable). This approach is slow and extremely expensive. It can take 15–18 years for a drug to go from a phase I study to standard of practice. Trastuzumab, the antibody directed against the Her2 oncoprotein, is an excellent example, where the time to get the drug tested and into common practice was 18 years. The FDA is extremely interested in new and more efficient approaches to trial design and has encouraged the use of adaptive designs [22]. I-SPY 2 is an adaptive study in high-risk early-stage breast cancers that is designed to accelerate the process of finding the right drugs for the right patients at the right time. It is focused on testing drugs at an earlier stage of the disease.

6.4.1 I-SPY 2 Description

I-SPY 2 is a multi-site, adaptive platform trial that has evolved into a platform for translational research. The goals are to drive drug development to the early-stage, high-risk setting, where women's lives can be saved, and to rapidly learn which combinations of agents have the best chance of curing patients.

Breast cancer is heterogeneous and ranges from indolent disease to very aggressive disease. As well, the timing of risk for recurrence can range from 3–5 years to 15–20 years. High-risk, fast-growing breast cancers can be characterized molecularly [23], and the risk for recurrence is largely in the first 5 years after diagnosis. These are the types of tumors that benefit from chemotherapy. Also, these are the tumors that, if you change the order of therapy and start with systemic therapy first (neoadjuvant chemotherapy), you can assess response to therapy by the time a patient has definitive surgical therapy. This simple change in the order of therapy does not change the outcome, but does allow much more efficient learning about what is working and for whom. And for these types of cancers, the degree of residual tumor is highly prognostic. Over the last 20 years, we and others have worked to establish an early endpoint, complete pathologic response (and residual cancer burden) [24], to predict 3-year event-free and distant recurrence-free survival [25].

There are several key innovations that have led to the success of this trial platform. One is the use of early endpoints to accelerate learning and changing the order of therapy to enable an early readout (pathology to measure response to therapy) in the course of care. The other is the use of biomarker and imaging guidance and the ability to identify the category of patients at risk for early recurrence for inclusion in the study. The third is the use of a platform, instead of a trial for every drug tested. The trial is designed pragmatically to test, in parallel, a number of agents/combinations so that the trial can become an engine for learning. Drugs come in and out of the trial by protocol amendment, saving significant time compared to writing and reviewing a new protocol for each new drug. The fourth is the use of new tools for real-time data capture. A longitudinal Bayesian adaptive model is used to predict "graduation" of agents that reach a threshold of an 85% predicted probability of success in a confirmatory 1:1 randomized neoadjuvant therapy trial of 300 patients. The purpose of the graduation threshold is to identify those agents with a big signal so we can focus on those combinations most likely to improve outcomes for patients. I-SPY 2 is a phase II trial designed to identify what is most likely to succeed in a focused subsequent phase III trial and avoid one of the biggest problems in oncology trials, where 70% of phase III trials fail. Finally, the trial is collaborative by design and has included the FDA, pharmaceutical and biotechnology companies, academic centers, and patients at the table from the inception. The trial design leverages a pre-competitive framework to align incentives and drive efficiency [26–28].

Several agents have graduated, and one of the most striking results was the near tripling of pathological complete response (pCR) with pembrolizumab added to Taxol, which was recently confirmed in a follow-on phase III trial [29, 30]. To date, 18 drugs and combinations have entered into the trial over the last decade, with many more in the pipeline [31–33]. There are currently 20 major academic sites in the network, with 4 more, including community cancer center networks, slated to enter in 2020. Over 20 companies participate. When the trial started, it was the first trial to bring multiple pharma companies into the same trial.

6.4.2 *Future of Adaptive Trials in the Neoadjuvant Setting*

Early endpoints are increasingly being accepted by both regulators and clinicians because they provide prognostic information that matters to patients [34]. The impact of achieving a pCR is very significant, and trials and care have evolved to adapt care based on response to therapy. This has led to a change in clinical practice, making it extremely difficult to conduct a standard randomized controlled trial. Much like with the HIV epidemic, patients and their oncologists know that their outcome will be better if they can achieve a complete response. Furthermore, a number of trials now demonstrate that targeted therapy can increase the chance of survival for those with substantial residual disease [31, 35].

These changes in practice are good for patients. However, it requires creativity in thinking about how to address clinical trials in the neoadjuvant setting. If you randomize women to standard of care or even an experimental drug plus standard of care and they do not get a great response, they will go on to get additional agents after surgery. That makes long-term endpoints extremely difficult to compare. Early endpoints are the key to advancing turns in knowledge. But once early endpoints are established, then clinicians and patients will not simply accept a poor outcome but will and should look for opportunities to further improve their chance of a good outcome. The implication is that our science and clinical trial designs will have to change.

Through a program project grant (P01CA210961) and the support of the Quantum Leap Healthcare Collaborative (trial sponsor), we are designing the next generation of I-SPY 2 with the explicit 5-year goal of getting 90% of patients to a pCR and a distant recurrence-free survival of 92% or higher. "I-SPY 2 Plus" will establish a new paradigm for clinical trials by encouraging the escalation/de-escalation of treatment depending upon an individual's response. I-SPY 2 Plus will leverage the Bayesian adaptive, biomarker-driven approach of the current I-SPY 2, combined with the "Sequential Multiple Assignment Randomized Trial" or "SMART" trial model that facilitates multiple randomizations within the same trial. In this hybrid model, patients who fail to respond to the therapy for which they have been initially randomized may be subsequently randomized to a second (and, in some cases, third) biologically targeted therapy as a "second chance" to achieve pCR. A further innovation will feature an additional, confirmatory arm that serves as a seamless transition from phase II to phase III development—a "Regulatory Evidence Generation" arm—designed to establish a more efficient means of gathering the evidence required for regulatory approval of an agent/combination. As I-SPY 2 Plus prepares for its launch transition in 2020, a number of supporting innovations in statistics and data acquisition and management are currently underway.

It is often true that trials are designed in a way that maximizes their chance of success, as journals and investigators are most interested in positive results. However, that is really not always the best chance to learn. In the era of personalized medicine, we should not be looking for large trials with a small benefit and a p value that is significant only because a large number of subjects are included. We should

be looking to find the interventions that have the greatest impact on specific subtypes of disease, and we should be designing trials to escalate interventions if risk is high and response to therapy is poor, while de-escalating therapy if risk is low and response is excellent. As well, we should be designing trials to enable rapid learning, such as adaptive (Bayesian) trials, which are more efficient. If such trials are designed to exploit the range of biology and cross disease signatures, what emerges can inform the field about where interventions are most likely to succeed. If there are early endpoints that are good surrogates of long-term outcomes, then we can design trials efficiently to get the most patients to the best early endpoints as quickly as possible.

The I-SPY trial highlights the need to develop pragmatic solutions to adapt to learning and change in the field and to enable rapid learning, even as new therapies emerge and standard of care evolves. It is critical that we design trials that can succeed and lead the way in an ever-changing environment.

6.5 Conclusions

There are many ways to learn, and randomized trials are one of the tools that should be employed to help advance the field. However, there are limitations to trials that have tight inclusion and exclusion criteria, and often, the findings from these trials may not apply to a broad population. In addition, there is inherent bias in a trial that requires randomization, as it only includes those who are willing to be randomized, or who specific physicians approach. Trials can be designed in pragmatic ways, to be more inclusive in the population, and also include populations that might otherwise be ignored. And there are ways to learn and modify guidelines in a pragmatic way without doing a trial and then use real-world evidence to generate the confirmatory evidence.

Our trials should be more patient-centered and designed to learn efficiently. Our trials should look more like our care, and our care should look more like pragmatic trials. The place we need to get to is one where learning naturally occurs in the course of care. Pragmatic trials are designed to help ensure our findings will be of greatest benefit to the patients we serve. But if we are to reach this goal, the pragmatic trial itself must continue to evolve.

References

1. Barnett AG, van der Pols JC, Dobson AJ. Regression to the mean: what it is and how to deal with it. Int J Epidemiol. 2004;34(1):215–20. https://doi.org/10.1093/ije/dyh299.
2. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. J Clin Epidemiol. 2009;62(5):499–505. https://doi.org/10.1016/j.jclinepi.2009.01.012.
3. Ford I, Norrie J. Pragmatic trials. N Engl J Med. 2016;375(5):454–63. https://doi.org/10.1056/NEJMra1510059.

4. Roland ME, Barin B, Huprikar S, et al. Survival in HIV-positive transplant recipients compared with transplant candidates and with HIV-negative controls. AIDS. 2016;30(3):435–44. https://doi.org/10.1097/QAD.0000000000000934.
5. Stock PG, Barin B, Murphy B, et al. Outcomes of kidney transplantation in HIV-infected recipients. N Engl J Med. 2010;363(21):2004–14. https://doi.org/10.1056/NEJMoa1001197.
6. Coffin CS, Stock PG, Dove LM, et al. Virologic and clinical outcomes of hepatitis B virus infection in HIV-HBV coinfected transplant recipients. Am J Transplant. 2010;10(5):1268–75. https://doi.org/10.1111/j.1600-6143.2010.03070.x.
7. Terrault NA, Roland ME, Schiano T, et al. Outcomes of liver transplant recipients with hepatitis C and human immunodeficiency virus coinfection. Liver Transpl. 2012;18(6):716–26. https://doi.org/10.1002/lt.23411.
8. Esserman LJ, Thompson IM, Reid B. Overdiagnosis and overtreatment in cancer. JAMA. 2013;310(8):797–8. https://doi.org/10.1001/jama.2013.108415.
9. Esserman LJ. The WISDOM study: breaking the deadlock in the breast cancer screening debate. npj Breast Cancer. 2017;3(1):34. https://doi.org/10.1038/s41523-017-0035-5.
10. Blue Cross Blue Shield Association. Blue cross and blue shield companies seek to personalize breast cancer screening through the WISDOM study. bcbs.com. https://www.bcbs.com/press-releases/blue-cross-and-blue-shield-companies-seek-personalize-breast-cancer-screening. Published 2 May 2018. Accessed 1 July 2019.
11. Rosenberg-Wohl S, Thygeson M, Stover Fiscalini A, et al. Private payer participation in coverage with evidence development: a case study. In: Health affairs. http://healthaffairs.org/blog/2017/03/14/private-payer-participation-in-coverage-with-evidence-development-a-case-study/. Published March 14, 2017. Accessed 4 Apr 2017.
12. eCQI Resource Center. Breast cancer screening eCQMs for 2018 performance period. ecqi.healthit.gov. https://ecqi.healthit.gov/ecqm/ep/2018/cms125v6. Published July 22, 2019. Accessed 10 Aug 2019.
13. United States Renal Data System. 2018 USRDS annual data report: epidemiology of kidney disease in the United States. usrds.org. https://www.usrds.org/adr.aspx. Published 2018. Accessed 1 Aug 2019.
14. Wong G, Webster AC. Cancer after kidney transplantation. In: Chapman JR, editor. Oxford textbook of clinical nephrology. Vol 1. Cancer after kidney transplantation. Oxford: Oxford University Press. https://doi.org/10.1093/med/9780199592548.003.0287.
15. Penn I. Cancers in renal transplant recipients. Adv Ren Replace Ther. 2000;7(2):147–56.
16. Knoll G, Cockfield S, Blydt-Hansen T, et al. Canadian Society of Transplantation: consensus guidelines on eligibility for kidney transplantation. CMAJ. 2005;173:S1–25. https://doi.org/10.1503/cmaj.1041588.
17. Mukhtar RA, Piper ML, Freise C, Veer LJV, Baehner FL, Esserman LJ. The novel application of genomic profiling assays to shorten inactive status for potential kidney transplant recipients with breast cancer. Am J Transplant. 2017;17(1):292–5. https://doi.org/10.1111/ajt.14003.
18. Wolf DM, Yau C, Wulfkuhle JD, et al. Integration of DNA repair deficiency and immune biomarkers to predict which early-stage triple-negative breast cancer patients are likely to respond to platinum-containing regimens vs. immunotherapy: the neoadjuvant I-SPY 2 trial. Chicago, IL; 2019.
19. Esserman LJ, Moore DH, Tsing PJ, et al. Biologic markers determine both the risk and the timing of recurrence in breast cancer. Breast Cancer Res Treat. 2011;129(2):607–16. https://doi.org/10.1007/s10549-011-1564-5.
20. Esserman LJ, Yau C, Thompson CK, et al. Use of molecular tools to identify patients with indolent breast cancers with ultralow risk over 2 decades. JAMA Oncol. 2017;3(11):1503–10. https://doi.org/10.1001/jamaoncol.2017.1261.
21. Mahajan R, Gupta K. Adaptive design clinical trials: methodology, challenges and prospect. Indian J Pharmacol. 2010;42(4):201–7. https://doi.org/10.4103/0253-7613.68417.
22. Woodcock J, LaVange LM. Master protocols to study multiple therapies, multiple diseases, or both. N Engl J Med. 2017;377(1):62–70. https://doi.org/10.1056/NEJMra1510062.

23. Sparano JA, Gray RJ, Makower DF, et al. Adjuvant chemotherapy guided by a 21-gene expression assay in breast cancer. N Engl J Med. 2018;379(2):111–21. https://doi.org/10.1056/NEJMoa1804710.
24. Symmans WF, Peintinger F, Hatzis C, et al. Measurement of residual breast cancer burden to predict survival after neoadjuvant chemotherapy. J Clin Oncol. 2007;25(28):4414–22. https://doi.org/10.1200/JCO.2007.10.6823.
25. Yee D, DeMichele A, Isaacs C, et al. Pathological complete response predicts event-free and distant disease-free survival in the I-SPY2 TRIAL (Abstract GS3-08). Cancer Res. 2018;78(4 Suppl) https://doi.org/10.1158/1538-7445.SABCS17-GS3-08.
26. Woodcock J. Precompetitive research: a new prescription for drug development? Clin Pharmacol Ther. 2010:521–3.
27. 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(1):97–100. https://doi.org/10.1038/clpt.2009.68.
28. Esserman LJ, Barker AD, Woodcock J, et al. A model for accelerating identification and regulatory approval of effective investigational agents. Cureus. 2012;4(12):e76. https://doi.org/10.7759/cureus.76.
29. Nanda R, Liu MC, Yau C, et al. Pembrolizumab plus standard neoadjuvant therapy for high-risk breast cancer (BC): results from I-SPY 2 (Abstract 506). J Clin Oncol. 2017;35(Suppl. 15) https://doi.org/10.1200/JCO.2017.35.15_suppl.506.
30. Merck. Merck's KEYTRUDA® (pembrolizumab) in combination with chemotherapy met primary endpoint of pathological complete response (pCR) in pivotal phase 3 KEYNOTE-522 trial in patients with triple-negative breast cancer (TNBC). merck.com. https://investors.merck.com/news/press-release-details/2019/Mercks-KEYTRUDA-pembrolizumab-in-Combination-with-Chemotherapy-Met-Primary-Endpoint-of-Pathological-Complete-Response-pCR-in-Pivotal-Phase-3-KEYNOTE-522-Trial-in-Patients-with-Triple-Negative-Breast-Cancer-TNBC/default.aspx. Published July 29, 2019. Accessed 1 Aug 2019.
31. Park JW, Liu MC, Yee D, et al. Adaptive randomization of neratinib in early breast cancer. N Engl J Med. 2016;375(1):11–22. https://doi.org/10.1056/NEJMoa1513750.
32. Rugo HS, Olopade OI, DeMichele A, et al. Adaptive randomization of veliparib–carboplatin treatment in breast cancer. N Engl J Med. 2016;375(1):23–34. https://doi.org/10.1056/NEJMoa1513749.
33. Esserman L, Hylton N, Asare S, et al. I-SPY2: unlocking the potential of the platform trial. In: Antonijevic Z, Beckman RA, editors. Platform trial designs in drug development umbrella trials and basket trials; 2018. p. 3–22.
34. Prowell TM, Pazdur R. Pathological complete response and accelerated drug approval in early breast cancer. N Engl J Med. 2012;366(26):2438–41. https://doi.org/10.1056/NEJMp1205737.
35. Masuda N, Lee S-J, Ohtani S, et al. Adjuvant capecitabine for breast cancer after preoperative chemotherapy. N Engl J Med. 2017;376(22):2147–59. https://doi.org/10.1056/NEJMoa1612645.

Chapter 7
Clinical Trials: Ensuring Quality and Standardization

Mihir M. Shah and Darren R. Carpizo

7.1 Introduction

The number of clinical trials in the United States and worldwide is increasing rapidly. ClinicalTrials.gov, which started in year 2000, registered a total of 2119 trials, whereas in 2019, that number is 309,531, which is nearly a 150-fold increase [1]. Of the 309,531 trials registered in 2019, 244,831 (79%) are interventional, with 25,857 - or approximately 11% - involving a surgical procedure. With this rise in the number of clinical trials, there is an increased need to ensure that human subjects research being conducted is done in a manner that is ethical and the data collected, analyzed, and reported are of high quality and transparent. To meet this need, organizations ranging from academic centers to large consortia that conduct clinical trials to the National Institutes of Health now have implemented training programs and defined standards for all professionals participating in the conduct of clinical trials. We will review some of these standards as well as other components of clinical trials to ensure quality and standardization.

M. M. Shah
Division of Surgical Oncology, Department of Surgery,
Emory University School, Atlanta, GA, USA

D. R. Carpizo (✉)
Division of Surgical Oncology, Department of Surgery,
University of Rochester, Rochester, NY, USA
e-mail: Darren_Carpizo@urmc.rochester.edu

T. M. Pawlik, J. A. Sosa (eds.), *Clinical Trials*, Success in Academic Surgery,
https://doi.org/10.1007/978-3-030-35488-6_7

7.2 Good Clinical Practice (GCP)

An international platform for establishing ethical and scientific quality standard for the following aspects of clinical trials involving human subjects—designing, conducting, recording, and reporting. This protects the rights, safety, and well-being of trial subjects [2].

The primary purpose of the International Council for Harmonisation (ICH) GCP guidelines is to enable mutual acceptance of clinical data by the following regulatory authorities—European Union, Japan, and United States—thereby offering a unified standard. Intention is to result in improvement in clinical trial quality and efficiency and producing reliable results while protecting human subjects [3].

7.3 The Collaborative Institutional Training Initiative (CITI)

The CITI program provides high-quality, peer-reviewed, web-based educational courses in research, ethics, regulatory oversight, responsible conduct of research, research administration, and other topics pertinent to the research enterprise [4]. Regularly designed and updated materials include:

- Enhancing the knowledge and professionalism of personnel conducting research.
- Education of members, administrators, and the leadership of the ethics committee.
- Promoting ethical research at organizations.

The CITI program was founded in March 2000. It now includes more than 20 subject areas (including biosafety and biosecurity, conflicts of interest, GCP, information privacy and security, and responsible conduct of research) [4]. A million learners access these materials annually at thousands of organizations. Many organizations, extending from universities to clinical research organizations and the National Institutes of Health (NIH), require CITI training certificates to conduct clinical trials. Effective as of 2017, the NIH requires GCP training to be re-certified every 3 years [5]. In addition, the CITI program not only offers GCP courses acceptable for the NIH policy, but also training on specific topics of interest to NIH-funded researchers [6].

7.4 Reporting of Adverse Events

7.4.1 Investigator Responsibility

The primary responsibility for adverse event (AE) identification, documentation, grading, and assignment of attribution is upon the clinical investigators, with ultimately the principal investigator being responsible for reporting it in a timely manner [7]. Any serious AE *must* be immediately reported to the sponsor by investigators [8].

7.4.2 Sponsor Responsibility

The sponsor should immediately inform the FDA and all participating investigators about any new significant AEs or risks with respect to investigational new drug (IND) or investigational device exemption (IDE). The following should be included in the annual report in addition to the summary of the previous year's clinical investigation [7]:

- Most frequent and most serious AEs.
- IND and IDE safety reports.
- Subjects who died (with the cause of death).
- Subjects who dropped out in association with AE (irrespective of its relation to the drug/device).

Grading of AE is critical and must be documented by medical personnel directly involved in the clinical care of protocol subjects [7]. Grading relates to the severity of AE for the purposes of reporting (Table 7.1).

Increasing evidence suggest that the severity and incidence of AEs are underestimated and underreported by physicians [9, 10]. Factors responsible for this include less attention to mild, subjective, or expected toxicities, less attention to toxicity in lieu of efficacy, increased patient volume, physician time constraints and limited resources for managing data, and unstructured/inadequate elicitation of all toxicities from the patient [9, 11]. In view of this, a standardized patient-centered method of AE reporting has been developed by the NCI—incorporation of this tool in clinical trials could help improve and complement physician reporting of AEs [12].

Table 7.1 Severity of adverse event

Grade	Description
0	*No AE* (within normal limits)
1	*Mild* Asymptomatic or mild symptoms, observations (clinical or diagnostic) only without any intervention
2	*Moderate* Minimal, local, or non-invasive intervention (e.g., packing, cautery) indicated, limiting age-appropriate instrumental activities of daily living (ADL)
3	*Severe* Medically significant but not immediately life-threatening, hospitalization or prolongation of hospitalization indicated, disabling, limiting self-care ADL
4	*Life-threatening consequences* Urgent intervention indicated
5	*Death* (related to AE)

Developed from: https://ctep.cancer.gov/protocolDevelopment/electronic_applications/docs/aeguidelines.pdf

7.5 Standard Operating Procedures (SOP)

A set of comprehensive instructions that define and standardize procedures in clinical trials are called SOP [13]. The main objective is to help the investigators and the research team stay compliant with the GCPs that govern the conduct of clinical research. SOP that are well-written and well-managed allow consistent execution of research-related activities in a standardized manner.

7.6 Case Report Form (CRF)

A document developed to record the required patient information to be reported to the sponsor in a clinical trial [14]. A well-designed CRF represents essential contents of the study protocol for optimal data collection [15]. This includes adverse events and serious adverse events. Figure 7.1 illustrates a typical serious adverse event form. The information that is recorded typically includes date and details of the event, comments from treating physician and the principal investigator, and classification of serious AE (Table 7.1), in addition to other details depicted in the form (Fig. 7.1). Nowadays, electronic CRFs (eCRFs) are typically used, as opposed to traditional paper CRF. eCRF is designed to facilitate capture of data with minimal errors, and regulatory authorities are willingly accepting submissions when validated systems are used to capture data electronically [16]. CRFs

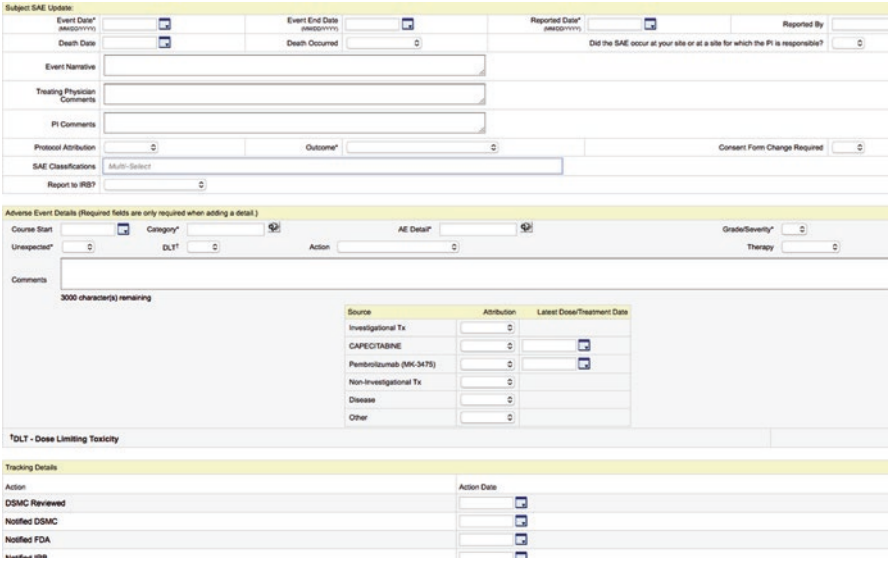


Fig. 7.1 Serious adverse event reporting form

are crucial to capture key data in order to allow a meaningful analysis. The guidelines to design a CRF are outside the scope of this chapter and are available in the literature [15].

7.7 Electronic Data Management Systems (EDMS)

EDMS are comprehensive standardized systems built through collaboration with prominent research organizations to streamline research operations. An example of such a system is OnCore.

OnCore is a premier application for managing clinical research data. OnCore not only has refined clinical research management functionality but also fully integrated patient registries, biospecimen management, billing compliance, and electronic data capture functionality. Over 50 leading clinical research institutions have chosen the OnCore system to manage and expand their portfolios of clinical trials [17].

Complion is another platform built by clinical researchers for clinical researchers. It is a cloud-based eRegulatory platform. It improves efficiency, compliance, and transparency for research sites and sponsors. It is established with the intention to institute the highest level of compliance with less amount of work, so the involved personnel can focus on what matters [18]. Complion is usually controlled by the regulatory department.

7.8 Data and Safety Monitoring Board (DSMB)

The primary roles of the DSMB include approval of the final protocol prior to enrollment and review of aspects of study progress periodically (patient enrollment, protocol compliance, data quality and completeness, adverse events, safety). In addition, the DSMB makes recommendations on continuation, modification, or termination of the trial [19].

The DSMB is essentially mandatory as it increases awareness of complications in clinical trials in order to ensure safety of the patients enrolled in the trial. The DSMB should consist of members that include at least one expert in the clinical aspects of the disease/patient population being studied, a biostatistician, and often ethicists [19].

The DSMB meetings are usually held at least annually and likely more often based on the requirement of the trial being monitored. Each meeting consists of an open session, closed session, and final executive session—the details of which can be found via this link: https://deainfo.nci.nih.gov/grantspolicies/datasafety.pdf.

For the safety of patients enrolled in the clinical trial and for the integrity of the data, the DSMB must have access to unmasked data. Some biostatisticians are of the opinion that the DSMB should be unmasked to treatment identity beginning from

the initial data review [20]. As far as the request of the DSMB for unmasking the trial data is honored, it is acceptable if the awardee leading the trial desires to keep some data masked.

7.9 Consolidated Standards of Reporting Trials (CONSORT)

The purpose of development of the CONSORT statement was to help authors report randomized controlled trials [21]. The statement offers a minimum standard set of items for reporting of clinical trials. In addition to improving quality of reporting in medical journals [22–25], it has been officially endorsed by hundreds of journals, including high-impact journals and prominent editorial groups [26]. It is advisable that the principal investigator should review the CONSORT statement while designing a clinical trial, and especially the checklist and the flowchart. Additional details can be viewed at the following link, where the updated guidelines of the 2010 CONSORT statement are published: https://www.bmj.com/content/340/bmj.c332.

7.10 Clinical Research Organization (CRO)

Multi-site clinical trial success depends on achieving quality across participating sites. Individual sites may not have the policies and procedures to accomplish this. Organizations like cooperative groups, consortia, and networks fill this need. These organizations have various attributes, such as commercial versus non-profit, being associated with a foundation or not, and specific geographical participation. Some provide comprehensive CRO-type resources, and others may offer their services "cafeteria style", including the site management organization (SMO) model.

Many efficiencies are gained using such an organization. They typically have created the policies and procedures rooted in common standards across participating sites. The organization ensures that these common standards meet the local regulatory requirements and comply with site institutional requirements. Services are provided centrally, allowing the organization to collect, compile, and confirm necessary documents and data per the protocol. Often, a Clinical Trial Management System (CTMS) is used. Various site regulatory documents are stored related to the particular study. Central processing of Adverse Events and Outside Safety Reports is another example of the way this kind of organization assures compliance across sites.

Consortia and similar organizations provide oversight to multi-site studies much like how an institution's clinical trial office oversees their own single-site study. Consortia can efficiently manage adherence to GCP across multiple sites with a systematic approach. Standardization is built into the study-specific database, usually within a Clinical Trial Management System, and automated processes assure

quality throughout the lifetime of the study [27]. These tools are used according to standard operating procedures written for the compliance needed in clinical trials.

Consortia incorporate language to operationalize the multi-site study within the protocol. Protocol-specific forms (eCRFs) are created for all sites to use, permitting standardized data collection. Site staff are trained on eCRF guidelines. Queries on data entry check pending and acceptable format. Queries are scheduled for the needed frequency, and output is sent to the appropriate staff within the consortia and at participating sites for resolution. This remote monitoring ensures accurate data is available for accrual and safety tracking. Alerts are programmed to detect stopping rule criteria and safety reporting triggers [28]. Consortia use processes like these to work with the site staff as a team ensuring quality [29].

7.11 Clinical Research Professionals (CRPs)

CRPs work in a wide range of settings and variety of organizations. Most centers require their CRPs to be certified in their roles via various examinations provided by organizations such as the Association of Clinical Research Professionals (ACRP) or the Society of Clinical Research Associates (SOCRA). Certification by such organizations attests to a standard of knowledge, education, and experience that is well recognized by the clinical research community. These standards promote recognition and continuing excellence in the ethical conduct of clinical trials [30]. Many institutions require their CRPs to obtain these certifications prior to their promotion to include greater responsibility and leadership opportunities.

7.12 Conclusion

Ensuring quality and standardization in today's clinical trials in surgery is no different than in pharmaceutical trials. It requires specialized training among all individuals participating in the study, not just the study investigators. There are now many online resources for obtaining such training and well-recognized organizations that provide online certifications. In the digital age, the key to creating high-quality data involves the building and use of high-quality software systems designed for the needs of each clinical trial and ensuring that the data entered is correct and updated regularly. In an age where more multi-center studies are needed, consortia in both academia and industry have created an even greater need to ensure data management systems are of high quality and meet the needs of investigators. These consortia have different practices, and it is imperative for the surgeon investigator to understand this and focus on this early in the trial's life to identify potential problems before they happen. Identifying problems early on will allow the changes to be made in the software to ensure data collection is of the highest quality for the remainder of the trials.

References

1. Clinical Trials—trends, charts and maps. https://clinicaltrials.gov/ct2/resources/trends.
2. E6(R2) Good clinical practice: integrated addendum to ICH E6(R1). U.S. Food and Drug Administration. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/e6r2-good-clinical-practice-integrated-addendum-ich-e6r1.
3. E6(R2) Good clinical practice: integrated addendum to E6(R1); International Council for Harmonisation; Guidance for Industry; Availability. https://www.regulations.gov/document?D=FDA-2018-D-0719-0001.
4. CITI program. https://about.citiprogram.org/en/mission-and-history/.
5. National Institutes of Health—policy and compliance. https://grants.nih.gov/policy/clinical-trials/good-clinical-training.htm.
6. Policy on good clinical practice training for NIH awardees involved in NIH-funded clinical trials. 2016. National Institutes of Health (NIH). https://grants.nih.gov/grants/guide/notice-files/NOT-OD-16-148.html.
7. NCI guidelines: adverse event reporting requirements. https://ctep.cancer.gov/protocolDevelopment/electronic_applications/docs/aeguidelines.pdf.
8. 21 CFR 312.64b, 21 CFR 812 (code of federal regulation title 21). U.S. Food and Drug Administration. https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/cfrsearch.cfm?fr=312.64.
9. Di Maio M, Gallo C, Leighl NB, Piccirillo MC, Daniele G, Nuzzo F, et al. Symptomatic toxicities experienced during anticancer treatment: agreement between patient and physician reporting in three randomized trials. J Clin Oncol. 2015;33(8):910–5. https://doi.org/10.1200/JCO.2014.57.9334.
10. Fromme EK, Eilers KM, Mori M, Hsieh YC, Beer TM. How accurate is clinician reporting of chemotherapy adverse effects? A comparison with patient-reported symptoms from the Quality-of-Life Questionnaire C30. J Clin Oncol. 2004;22(17):3485–90. https://doi.org/10.1200/JCO.2004.03.025.
11. Montemurro F, Mittica G, Cagnazzo C, Longo V, Berchialla P, Solinas G, et al. Self-evaluation of adjuvant chemotherapy-related adverse effects by patients with breast cancer. JAMA Oncol. 2016;2(4):445–52. https://doi.org/10.1001/jamaoncol.2015.4720.
12. Basch E, Reeve BB, Mitchell SA, Clauser SB, Minasian LM, Dueck AC, et al. Development of the National Cancer Institute's patient-reported outcomes version of the common terminology criteria for adverse events (PRO-CTCAE). J Natl Cancer Inst. 2014;106(9) https://doi.org/10.1093/jnci/dju244.
13. Sajdak R, Trembath L, Thomas KS. The importance of standard operating procedures in clinical trials. J Nucl Med Technol. 2013;41(3):231–3. https://doi.org/10.2967/jnmt.113.121467.
14. ICH guidance E6: good clinical practice: consolidated guideline. US HHS, US FDA, CDER, CBER. 1996. http://www.fda.gov/downloads/Drugs/Guidances/ucm073122.pdf. Accessed 11 Jun 2013.
15. Bellary S, Krishnankutty B, Latha MS. Basics of case report form designing in clinical research. Perspect Clin Res. 2014;5(4):159–66. https://doi.org/10.4103/2229-3485.140555.
16. Good clinical data management practices. Belgium: Society for Clinical Data Management; 2009. Accessed 17 Jun 2013. http://www.swaggerinfosolutions.pro/sridevi/New%20Folder/ICRI/GCDMP_complete_OCT2009.pdf.
17. OnCore enterprise research CTMS. Forte. https://forteresearch.com/enterprise-research-oncore/.
18. eRegulatory solution for high-performing clinical research sites. Complion. https://complion.com/.
19. Holubkov R, Casper TC, Dean JM, Anand KJ, Zimmerman J, Meert KL, et al. The role of the Data and Safety Monitoring Board in a clinical trial: the CRISIS study. Pediatr Crit Care Med. 2013;14(4):374–83. https://doi.org/10.1097/PCC.0b013e318274568c.
20. Piantadosi S. Clinical trials: a methodologic perspective. New York: John Wiley & Sons; 2005.

21. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. JAMA. 1996;276(8):637–9.
22. Moher D, Jones A, Lepage L, CONSORT Group. Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation. JAMA. 2001;285(15):1992–5.
23. Plint AC, Moher D, Morrison A, Schulz K, Altman DG, Hill C, et al. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. Med J Aust. 2006;185(5):263–7.
24. Turner L, Shamseer L, Altman DG, Weeks L, Peters J, Kober T, *et al.* Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals. Cochrane Database Syst Rev. 2012;11:MR000030. https://doi.org/10.1002/14651858.MR000030.pub2.
25. Devereaux PJ, Manns BJ, Ghali WA, Quan H, Guyatt GH. The reporting of methodological factors in randomized controlled trials and the association with a journal policy to promote adherence to the Consolidated Standards of Reporting Trials (CONSORT) checklist. Control Clin Trials. 2002;23(4):380–8.
26. Shamseer L, Hopewell S, Altman DG, Moher D, Schulz KF. Update on the endorsement of CONSORT by high impact factor journals: a survey of journal "Instructions to Authors" in 2014. Trials. 2016;17(1):301. https://doi.org/10.1186/s13063-016-1408-z.
27. 21 CFR 11. U.S. Food and Drug Administration. https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?CFRPart=11.
28. 21 CFR 312.32–33. U.S. Food and Drug Administration. https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRsearch.cfm?CFRPart=312.
29. 21 CFR 312.56. U.S. Food and Drug Administration. https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?fr=312.56.
30. Cuccia M. How to advance your career through clinical research certification. Forte. https://forteresearch.com/news/how-to-advance-your-career-through-clinical-research-certification/.

Chapter 8
Steps in Device and Drug Pathway Development: Clinical Trials, Similarities, and Differences

Timur P. Sarac

8.1 Preliminary Work Before Human Trials

8.1.1 The Idea and Documentation

New devices and drugs are developed from unmet needs. The range of the unmet need varies; for example, there may be inadequate therapeutic options for a malignancy, or there are no minimally invasive options to treat complex ascending aortic aneurysms. In order to protect your intellectual property, it is important to be as descriptive as you can and document your descriptions in dated files or notebooks.

8.1.2 The Patent

Obtaining a patent legally protects you from anyone else using your idea. The process of obtaining a patent can be time consuming and very expensive. The average United States patent costs approximately $20,000. Extending this to the European Union, Asian, and South American countries in conglomerate can cost over $100,000. The initial process begins with confirming whether or not there is an existing patent with the same or similar idea. Prior to retaining a patent attorney, one can do the search on their own through the United States Patent and Trademark Office (USPTO.gov) or through Google Patents (patents.google.com). Do not be discouraged if you happen to find "prior art," as your idea may have a different slant. Once you are convinced of the uniqueness of your idea or invention, hiring a patent

T. P. Sarac (✉)
The Ohio State University School of Medicine, Wexner Medical Center, The Ross Heart Hospital, Columbus, OH, USA
e-mail: Timur.Sarac@osumc.edu

T. M. Pawlik, J. A. Sosa (eds.), *Clinical Trials*, Success in Academic Surgery,
https://doi.org/10.1007/978-3-030-35488-6_8

attorney is the next step. The patent attorney will also research the idea, and if there is agreement that there is no prior art, then the attorney will usually have you file a provisional patent application. This provisional application sets the date and time for patent protection, but not the final content [1]. A provisional application includes a description and drawing(s) of an invention. Drawings may be required for understanding the subject matter sought to be patented but not formal patent claims, inventors' oaths, or declarations [2]. The USPTO does not conduct a formal review of provisional patents, and therefore a provisional application will never become a patent on its own. However, the provisional patent application can be subsequently converted into a non-provisional "utility patent" application by the applicant, at which point the application is examined as a utility application [3]. The provisional application is also not "published," but becomes a part of the record, along with any later non-provisional application file that references it, and thus becomes "public" upon issuance of a patent claiming its priority benefit. This allows you 1 year to fine-tune the description and argument which will be put into writing. This final "utility patent" consists of a thorough description of the invention, illustrations, and lastly "claims," which are the specifics of the idea that are unique and protected from anyone else commercializing the invention.

Once the patent is filed, you may not hear back from the patent and trademark office for 18–24 months. The patent is assigned to an examiner who is part of a special arts unit in the field of the invention. It may take 18 months for the examiner to even pick up the patent. Once the examiner reviews the patent, he/she searches again for prior art both in the United States and abroad. When this is complete, they will issue an office action, which outlines the examiner's findings. It is very unusual for them to issue the patent immediately. More likely, they will reject several aspects of the patent, and you will then be given an opportunity to respond. After several go rounds, several more months, and many bills, a conclusion is finally reached. Hopefully, your idea will be unique enough to get a patent, which starts with a USPTO issuance of what is termed as "Notice of Allowance" for the unique claims, which results in an issued patent. There are several more fees for this. The patent's life is 20 years, and over the next 20 years, you are required to pay maintenance fees for the patent. In addition, it is not unusual for the patent examiner to say there are more than one invention within the patent and ask you to separate out the two, with the second idea being filed as a "divisional" patent.

An important follow-up is that while you file the original patent and continue to work on your invention, new and additional material may come to light. In these circumstances, you can add to the patent by filing either a "continuation" or a "continuation in part (CIP)" with the United States Patent and Trademark Office. A CIP application permits a patent applicant to add new subject matter to the existing disclosure of the parent application, while retaining the priority date for claims based on the original disclosure [4]. The claims of the CIP can be directed to the new subject matter, the old subject matter, or a combination of the two. In contradistinction, "continuation application" inherits the parent application's priority date—but it is limited to the parent application's disclosure. New developments since the original patent filing cannot be described in a continuation application, as new matter

cannot be added to a continuation application. Instead, a continuation allows an inventor to add new claims to a parent application, as long as the application has not been approved or abandoned. With a continuation application, an inventor may increase the scope of his application without having to file an entirely new application, avoiding losing the original filing date. An important point for continuation applications is that they may only edit the claims.

8.1.3 Prototypes, Studies, Nuts, and Bolts

Concurrent to filing the patent, developing a drug or device requires a basic prototype or proof of concept. This leads to justification for continuing or altering your current path. It is important to keep extremely accurate records, as the Food and Drug Administration (FDA) requires you to maintain a design history file [5].

It is helpful to make computer animated designs and perform finite element analysis (FEA) prior to construction of prototypes for both medical devices and drugs. This allows for computer simulation of areas of fatigue. Traditionally, this is necessary for new drugs designed from X-ray crystallography [6]. The FDA will want to review this during their safety analysis. Next comes the prototype phase, which usually has the intent of proof of concept. Following satisfactory prototype construction, there are several required tests necessary before you can proceed with human trials. Some examples include mechanical strength testing, accelerated fatigue analysis, bioburden testing, and acute and chronic animal studies. For medical devices and drugs, the FDA has published Office of Device Evaluation "Guidance Documents" which provide an outline to follow and these can be found on their website [7]. It is important to note that to proceed with a clinical trial for a new implantable medical device or drug, the design of the device or drug cannot change once you initiate clinical trials. This is called "design freeze." Any change may precipitate the need for the inventor to repeat several of the preclinical steps, including new animal studies.

The cost to bring a new medical device and new drug to market is incredibly expensive. The preclinical work alone can surpass $10,000,000. The clinical trials themselves are even more expensive, because the regulatory pathways are stringent and require strict auditing for extended periods. Funding can come in the way of grants (e.g., SBIR/STTR), venture capital, or a major technology company venture arm. Contract research organizations (CROs) evolved to manage the process including the trials, and their industry itself has grown to several billion dollars. These companies provide support to the pharmaceutical, biotechnology, and medical device industries in the form of research services outsourced on a contract basis. A CRO may provide services such as biopharmaceutical development, commercialization, preclinical research, clinical research, clinical trial management, X-ray core lab, and pharmacovigilance for adverse outcomes. Many CROs specifically provide clinical-study/clinical-trial support for drugs and/or medical devices and range from large, international full-service organizations to small, niche specialty groups.

CROs that specialize in clinical-trial services can offer their clients the expertise of moving a new drug or device from its conception to FDA/EMA marketing approval, without the drug sponsor having to maintain a staff for these services [8].

An important component of all trials involving new medical device and drugs is the establishment of a Data Safety and Monitoring Board (DSMB), an independent group of experts who monitor patient safety and treatment efficacy data while a clinical trial is ongoing [9]. The primary mandate of the DSMB is to protect patient safety. If a serious adverse event (SAE) occurs more commonly in the experimental arm compared to the control arm, then the DSMB would strongly consider termination of the study. This evaluation has to be made in consideration of risk/benefit of the new therapy. In many cases, the experimental arm could cause serious adverse events (e.g., chemotherapy), but the resulting improvement in survival may outweigh the adverse events. If the experimental arm is proven to be undeniably superior to the control arm, the DSMB may recommend termination of the trial. This would allow the company sponsoring the trial to get regulatory approval earlier and to allow the superior treatment to get to the patient population earlier. This is usually an uncommon event, as the statistical evidence needs to be very high through prescribed power analyses done prior to the study initiation. Also, there may be other reasons to continue the study, such as collecting more long-term safety data. Futility is not as widely recognized as safety and benefit, but it actually can be the most common reason to stop a trial. For example, suppose a trial is one-half completed, but the experimental arm and the control arm have nearly identical results. It's likely in no one's interest to have this trial continue. In this circumstance, it is extremely unlikely that the trial would continue to its normal end, as it would be unlikely that there would be statistical evidence needed to convince a regulatory agency to approve the treatment [10]. The company sponsoring the study could save money for other projects by abandoning this trial. Also, current and potential trial participants could be freed to take other treatments rather than continue this experimental treatment which is unlikely to benefit them.

8.2 Human Trials: Medical Devices

The Medical Device Amendments of 1976 to the Federal Food, Drug, and Cosmetic Act established three regulatory classes for medical devices. The three classes are based on the degree of control necessary to assure the various types of devices are safe and effective [11]. Class I—The device presents minimal potential for harm to the patient. Forty-seven percent of all medical devices fall into this category, and 95% are exempt from the regulatory process. If a device falls into a generic category of exempted class I devices, a premarket approval (PMA) application (PMA) and FDA clearance are not required (see below) before marketing the device in the United States. However, the manufacturer is required to register their establishment and list their generic product with FDA. Some examples may include band aids, enemas, manual stethoscopes, and mercury thermometers. Class II—These pose

some risk, but in general the risks are very minor. Some examples include powered wheelchairs, crutches, and some over-the-counter testing kits. Forty-three percent of medical devices fall under this category. Class III—These devices are deemed necessary to sustain and support life and limb. They are commonly implanted or present potential risk of illness or injury. Examples of class III devices include implantable pacemakers and stents. Ten percent of medical devices fall under this category. These devices require a premarket approval (PMA) from the FDA [12]. Manufacturers must submit a premarket approval (PMA) application to the FDA if they wish to market any new products that differ in design or contain new materials from products already on the market. A PMA submission must provide valid scientific evidence collected from preclinical and human clinical trials showing the device is safe and effective for its intended use. If the device you are researching is life sustaining or presents any potential of unreasonable risk of illness or injury, you should search FDA's premarket approval (PMA) releasable database [13]. All United States and many foreign countries' clinical trials are required to be registered on ClinicalTrials.gov.

The pathway to market for a PMA may follow one of four pathway applications. The first is a 510(k) application; the second is an investigational device exemption (IDE); the third is a humanitarian device exemption (HDE); and the fourth is an investigational new drug (IND). The following will describe each of these:

8.2.1 510(k)

Once the preclinical data are assembled, the next step in the PMA process may be to perform a human clinical safety and efficacy trial, which will be described below. However, on occasion, a simpler pathway is available, called a "premarket notification 510(k)" submission. A 510(k) is a premarket submission made to the FDA to demonstrate that the device to be marketed is *at least* as safe and effective as another predicate device deemed substantially equivalent. Only a small percentage of 510(k) s require clinical data to support the application. Another situation may be that investigational use also includes clinical evaluation of modifications of a current device or new intended uses of legally marketed devices. In these circumstances, while preclinical work is needed, it is not subject to market approval like a (PMA 21CR 807.92(a)(3)) [14]. Only a small percentage of 510(k)s require clinical data to support a marketing clearance by the FDA. Under a 510(k) application, before a manufacturer can market a medical device in the United States, they must demonstrate to the FDA's satisfaction that it is substantially equivalent (as safe and effective) to a device already on the market. If FDA rules the device "substantially equivalent," the manufacturer can market the device. If the device you are researching has been in commercial distribution before 1976 or is substantially equivalent to a device already on the market, you should search FDA's 510(k) releasable database [15]. After 1976, 510(k) is a premarket submission made to the FDA for a legally marketed device (21 CFR 807.92(a)(3)) [14] that is not subject to PMA. Submitters

must compare their device to one or more similar legally marketed devices and make and support their substantial equivalence claims through preclinical data. The legally marketed device to which equivalence is drawn is commonly known as the "predicate." Although devices recently cleared under 510(k) are often selected as the predicate to which equivalence is claimed, any legally marketed device may be used as a predicate.

A device is substantially equivalent if it has the same *intended use* as the predicate. Additionally, it must have the same technological characteristics as the predicate. Another option is that it has the same intended use as the predicate, but it has different technological characteristics and does not raise different questions of safety and effectiveness, and the information submitted to FDA demonstrates that the device is at least as safe and effective as the legally marketed device. A claim of substantial equivalence does not mean the new and predicate devices must be identical. Substantial equivalence is established with respect to intended use, design, energy used or delivered, materials, chemical composition, manufacturing process, performance, safety, effectiveness, labeling, biocompatibility, standards, and other characteristics.

All manufacturers (including specification developers) of class II and III devices and select class I devices are required to follow design controls (21 CFR 820.30) during the development of their device [16]. The holder of a 510(k) must have design control documentation available for FDA review during a site inspection. In addition, any changes to the device specifications or manufacturing processes must be made in accordance with the quality system regulation (21 CFR 820) and may be subject to a new 510(k). Additional information is found on the webpage: "Is a new 510(k) required for a modification to the device?" Of note, the FDA does not perform 510(k) pre-clearance facility inspections. The submitter may market the device immediately after 510(k) clearance is granted. The manufacturer should be prepared for an FDA quality system (21 CFR 820) inspection at any time after 510(k) clearance [17].

8.2.2 Investigational Device Exemption (IDE)

An investigational device exemption (IDE) [18] allows an investigational device that is the subject of a clinical study to be used in order to collect safety and effectiveness data required to support a PMA application. All clinical evaluations of investigational devices, unless exempt, must have an approved IDE by the FDA before the study is initiated. The FDA does not disclose the existence of an IDE because the information is considered confidential. The IDE application and process can be long and arduous. Fortunately, the FDA has a mechanism to escort the medical device developer through the process by doing a "pre-IDE meeting" [19]. An IDE limits the distribution of an investigational device only to the sites identified in the IDE application. In addition to FDA requirements, clinical studies of devices are also monitored by institutional review boards (IRBs) located at hospitals or other facilities where the clinical studies are conducted. The purpose of an IRB's review is to assure

ethical principles are in place for patient selection criteria and that adequate informed consent information is provided to highlight the risks to patients. The IRB acts as the FDA's surrogate to oversee the protection of human subjects who participate in the clinical studies. The initial risk determination of a clinical study and/or device is made by an IRB in most cases. The IRB determines if a device/clinical study is significant or non-significant risk. The FDA can overrule any risk determination made by an IRB. If the IRB determines that a device/clinical study is a significant risk, the applicant must submit an IDE application to the FDA. The FDA must approve the application prior to the applicant enrolling patients in the clinical study. If the IRB determines that the clinical study/device is non-significant risk, the applicant can enroll patients without submitting an IDE application to the FDA. The clinical study will be monitored by the IRB under the abbreviated requirements of the IDE regulations in *21 CFR 812.2*(b) [20]. Typical requirements include informed consent from all patients, labeling for investigational use only, strict and close monitoring of the study, and meticulous records and reports. CROs can do much of this.

An approved IDE permits a device to be shipped lawfully for the purpose of conducting investigations of the device without complying with other requirements of the Food, Drug, and Cosmetic Act that would apply to devices in commercial distribution. Therefore, the sponsors are not required to submit a PMA or premarket notification, register their establishment, or list the device while the device is under investigation. Sponsors of IDEs are also exempt from the quality system (QS) regulation except for the requirements for design control [21]. There are no preprinted forms for an IDE application; however, an IDE application must include certain required information. For example, the sponsor must demonstrate that there is no reason to believe that the risks to human subjects from the proposed investigation outweigh the anticipated benefits to subjects. Others include demonstrating that the importance of the knowledge to be gained in the study is scientifically sound and that there is reason to believe that the device as proposed for use will be efficacious.

Once an IDE application is submitted, the FDA has 90 days to approve. However, in the 90-day period, the clock can stop, as the FDA may send back several questions that they want you to address before they allow you to proceed. An IDE study, like an IND below, takes form in clinical trial phases. Medical device clinical trials are different from drug trials in that only patients with the condition which the device is designed to treat are involved. They are traditionally comprised of three different types of studies [22]:

1. Feasibility study—Feasibility studies are the first human studies conducted in device development. They are used to establish preliminary safety and effectiveness of the device. They also set the study design for the next stage of the trial, the pivotal study. The emphasis on these trials is placed on safety, and the study numbers are small.
2. Pivotal study—Pivotal studies are also performed to demonstrate that the device is safe and effective for a specific use within a defined patient population. The patient numbers are significantly higher, and the results of a pivotal study are used to gain final regulatory approval to market the device through the PMA.

3. Post-market study—These are run either as a condition of approval in addition to meet a business objective. Post-market studies are similar to phase IV clinical drug trials, in that the goal is to better understand long-term effectiveness of the device and potential adverse events associated with the use of the device over extended periods of time.

8.2.3 Humanitarian Device Exemption (HDE)

Humanitarian device exemptions are available for devices where the rarity of the disease precludes enough patient numbers to satisfy the safety and effectiveness of scrutinized clinical IDEs. The Humanitarian Use Device (HUD) Program designates medical devices that are intended to benefit patients in the treatment or diagnosis of a disease or condition that affects or is manifested in not more than 8000 individuals in the United States per year as eligible for humanitarian device exemption (HDE). The concept came from the Orphan Drug Act (ODA) of 1984, where a rare disease is defined as a condition that affects fewer than 200,000 people in the United States. Currently, in the United States, only a portion of the 7000 known rare diseases have approved treatments. As a result, it has been difficult to gather enough clinical evidence to meet the FDA standard of reasonable assurance of safety and effectiveness. In order to address this challenge, Congress included a provision in the Safe Medical Devices Act of 1990 to create a new regulatory pathway for products intended for diseases or conditions that affect small (rare) populations. This is the HDE Program. An HDE is an approval process provided by the United States FDA allowing a medical device to be marketed without requiring evidence of effectiveness, but does maintain the safety component. The FDA calls a device approved in this manner a "humanitarian use device" (HUD) [23]. Under section 520(m)(6)(A)(i) of the FD&C Act, an HUD is only eligible to be sold for profit after receiving approval if the device is intended for the treatment or diagnosis of a disease or condition that falls under one of two categories. First, it occurs in pediatric patients or in a pediatric subpopulation, and such device is labeled for use in pediatric patients or in a pediatric subpopulation in which the disease or condition occurs. Second, it occurs in adult patients in such low numbers that the traditional pathway for development of the device for such patients is impossible, highly impracticable, or unsafe. The number of HDE devices that may be sold for profit is limited to a quantity known as the annual distribution number (ADN). If the FDA determines that an HDE holder is eligible to sell the device for profit, the FDA will determine the ADN and notify the HDE holder. The ADN is calculated by taking the number of devices reasonably necessary to treat or diagnose an individual per year and multiplying it by 8000. For example, if the typical course of treatment using an HDE device, in accordance with its intended use, requires the use of two devices per patient per year, then the ADN for that HDE device would be 16,000 (i.e., 2×8000). If the number of devices distributed in a year exceeds the ADN, the sponsor can continue to sell the device but cannot earn a profit for the remainder of the year.

8.3 Human Trials: New Drugs

8.3.1 Investigational New Drug (IND)

The FDA's role in the development of a new drug begins when the drug's sponsor (usually the manufacturer), having screened the new molecule for pharmacological activity and acute toxicity potential in animals, wants to test its diagnostic or therapeutic potential in humans. At that point, the molecule changes in legal status under the Federal Food, Drug, and Cosmetic Act and becomes a new drug subject to specific requirements of the drug regulatory system. This takes form as an IND application. Current federal law requires that all drugs are required to be approved by specified applications before they are transported or distributed across state lines. Because a sponsor will probably want to ship the investigational drug to clinical investigators in many states, it must seek an exemption from that legal requirement. The IND is the means through which the sponsor technically obtains this exemption from the FDA [24].

INDs are similar to IDEs in that they are long arduous processes which take many years to navigate both preclinical and clinical work. The FDA recognizes this and therefore provides guidance documents to outline the required steps. The guidance documents represent the FDA's current thinking on a particular subject [24]. These documents provide FDA review staff and applicants/sponsors with guidelines to the processing, content, and evaluation/approval of applications and also to the design, production, manufacturing, and testing of regulated products. They also establish policies intended to achieve consistency in the FDA's regulatory approach and establish inspection and enforcement procedures. For both devices and drugs, guidance documents are not regulations or laws, so they are not enforceable, either through administrative actions or through the courts. Alternative approaches to the guidance documents may be used if it satisfies the requirements and are negotiated with the FDA. A list of the specified areas and documents can be found on their website [25, 26].

There are two IND categories, which are commercial and research. These are further divided into three types, which subsequently will be described. First, an *Investigator IND* is submitted by a physician who both initiates and conducts an investigation and under whose immediate direction the investigational drug is administered or dispensed. A physician might submit a research IND to propose studying an unapproved drug or an approved product for a new indication or in a new patient population. The second is an *Emergency Use IND*, which allows the FDA to authorize use of an experimental drug in an emergency situation that does not allow time for submission of an IND in accordance with specific outlined policies (21CFR, *Section 312.23* or *Section 312.20*) [27]. It is also used for patients who do not meet the criteria of an existing study protocol, or if an approved study protocol does not exist. The third is a *Treatment IND*, which is submitted for experimental drugs showing promise in clinical testing for serious or immediately life-threatening conditions while the final clinical work is conducted and the FDA review takes place.

During a new drug's early preclinical development, the sponsor's primary goal is to determine if the product is reasonably safe for initial use in humans and if the compound exhibits pharmacological activity that justifies commercial development. When a product is identified as a viable candidate for further development, the sponsor then focuses on collecting the data and information necessary to establish that the product will not expose humans to unreasonable risks when used in limited, early-stage clinical studies. Given the complexity of the process, again similar to IDEs, the FDA provides a means to facilitate a streamlined process through a pre-IND meeting [28]. Several meetings may be required, and the FDA can give guidance as to whether the application is adequate or whether additional preclinical work is necessary. The review divisions are organized generally along therapeutic classes and can each be contacted from the online site. Additionally, any biologic product may fall under an IND even if it is considered a device, and guidance should be sought as to which pathway is required [29].

The IND application must contain information in the following three broad areas, which are very similar to medical device requirements:

1. Animal pharmacology and toxicology studies—This is preclinical data to permit an assessment as to whether the product is reasonably safe for initial testing in humans. Also included are any previous experiences with the drug in humans (often foreign use).
2. Manufacturing information—This is information pertaining to the composition, manufacturer, stability, and controls used for manufacturing the drug substance and the drug product. This information is assessed to ensure that the company can adequately produce and supply consistent batches of the drug. This is similar to QA procedures for medical devices.
3. Clinical protocols and investigator information—This provides detailed protocols for proposed clinical studies to assess whether the initial-phase trials will expose subjects to unnecessary risks. In this process, a thorough evaluation is done on the trial physician's capabilities. Finally, this contains commitments to obtain informed consent from the research subjects, to obtain review of the study by an institutional review board (IRB), and to adhere to the investigational new drug regulations.

Once the IND is submitted, the sponsor must wait 30 calendar days before initiating any clinical trials. During this time, the FDA has an opportunity to review the IND for safety to assure that research subjects will not be subjected to unreasonable risk. However, if the FDA has questions, the 30-day period stops until the questions are answered to the FDA's satisfaction.

Similar to IDE Applications, INDs set the investigator up for clinical trials. New drug clinical trials for drugs are divided into four phases:

- Phase I—The drug is tested on a small group of healthy individuals. This phase is used to determine the appropriate dosing, how humans react to the drug, and possible side effects. This phase is different than medical devices, as implants are not done on healthy patients.

- Phase II—Here, the drug is administered to a larger group of people, usually divided into two groups: one which receives the experimental drug and one which receives a placebo. This part of the study allows researchers to determine the relative safety and effectiveness of the drug. Again, this is different than medical devices, as there is no placebo group in medical device trials.
- Phase III—This phase involves testing the drug on a larger population (between several hundred and several thousand individuals) to confirm its effectiveness, its benefits, how it compares to other treatments, and possible adverse reactions. When phase III is complete, pharmaceutical companies can request FDA approval to introduce the drug to the market.
- Phase IV—Often referred to as post-marketing surveillance trials and similar to medical devices, phase IV studies are conducted after the drug has received permission from the FDA to be sold. In this phase, pharmaceutical companies can compare their drug to other drugs in the market and monitor the drug's long-term efficacy.

8.3.2 New Drug Accelerated Process: Orphan Drug Application [30]

In rare circumstances, the traditional pathway to getting a new drug approved is not feasible and a potentially significant therapeutic product is withheld. The FDA has another pathway to thoroughly evaluate these "orphan drugs," similar to an HDE for medical devices. The Office of Orphan Products Development's (OOPD) mission is to advance the evaluation and development of products (drugs, biologics, devices, or medical foods) that demonstrate promise for the diagnosis and/or treatment of rare diseases or conditions. The OOPD evaluates scientific and clinical data submissions from sponsors to identify and designate products as promising for rare diseases and to further advance scientific development of such promising medical products. The office also works on rare disease issues with the medical and research communities, professional organizations, academia, governmental agencies, industry, and rare disease patient groups [31].

OOPD provides incentives for sponsors to develop products for rare diseases. The program has successfully enabled the development and marketing of over 600 drugs and biologic products for rare diseases since 1983. In contrast, fewer than ten such products supported by industry came to the market between 1973 and 1983. The Orphan Grants Program has been used to bring more than 60 products to marketing approval. The Humanitarian Use Device Program has been the first step in approval of 70 humanitarian device exemption approvals. The Orphan Drug Designation Program provides orphan status to drugs and biologics which are defined as those intended for the safe and effective treatment, diagnosis, or prevention of rare diseases/disorders affecting either fewer than 200,000 people in the United States or more than 200,000 persons but they are not expected to recover the

Table 8.1 Comparison of new device and new drug pathways to approval

	IND	IDE	510(k)
	Preclinical		
Animal	Yes	Yes	Variable
Pharmacology and toxicology	Yes	Yes	Yes
Design control	Yes	Yes	Yes
Design history file	Yes	Yes	Yes
	Clinical		
PMA application	Yes	Yes	Variable
Pre-FDA meeting	Yes	Yes	Yes
Phases	4	3	Variable
Healthy volunteer testing	Yes	No	No
Placebo in pivotal trial	Yes	No	No
Rare diseases	Orphan drug	HDE	HDE
DSMB	Yes	Yes	No
ClinicalTrials.gov required	Yes	Yes	Yes

Grants Program [32]

costs of developing and marketing a treatment drug. Another component of this is the Rare Pediatric Disease Priority Review Voucher Program. This program permits a sponsor who receives an approval for a drug or biologic for a "rare pediatric disease" to qualify for a voucher that can be redeemed to receive a priority review of a subsequent marketing application for a different product. The OOPD administers three extramural grant programs: the Orphan Products Clinical Trials Grants Program which provides funding for clinical research that tests the safety and efficacy of drugs, biologics, medical devices, and medical foods in rare diseases or conditions, the Orphan Products Natural History Grants Program which supports studies that advance rare disease medical products development through characterization of the natural history of rare diseases and conditions, and the Pediatric Device Consortia (PDC) Grants Program which provides funding to develop nonprofit consortia to facilitate pediatric medical device development (Table 8.1).

References

1. Pressman D. Patent it yourself. Berkeley, CA: Nolo Press; 2006. p. 56. ISBN 1-4133-0516-4.
2. https://www.bitlaw.com/source/35usc/113.html.
3. Chisum D. Volume 4A, Chapter 14. Priority by foreign filing, § 14.02 historical development. In: Chisum D, editor. Chisum on patents. Seattle: Matthew Bender LexisNexis; 2010.
4. http://patentattorneycionca.com/PatentInformation/DivisionalPatent.
5. Harnack G. Mastering and managing the FDA maze: medical device overview. Milwaukee, American Society for Quality. .ISBN 9780873894555; 1999.
6. Giacovazzo C. Herbert Hauptman (1917–2011) Mathematician whose theories reveal the shapes of molecules from scattered X-rays. Nature. 2011;479:300.

7. https://www.fda.gov/medicaldevices/deviceregulationandguidance/guidancedocuments/default.htm.
8. https://www.statista.com/statistics/732804/top-clinical-research-organizations-by-revenue/.
9. Herson J. Data and safety monitoring committees in Clinical Trials. Boca Raton, FL: CRC; 2009. ISBN 978-1-4200-7037-8.
10. Snappin S, Chen MG, Jiang Q, Koutsoukos T. Assessment of futility in clinical trials. Pharm Stat. 2006;5(4):273–81.
11. https://www.fda.gov/medicaldevices/resourcesforyou/consumers/ucm142523.htm.
12. https://www.fda.gov/medicaldevices/deviceregulationandguidance/howtomarketyourdevice/premarketsubmissions/premarketapprovalpma/.
13. https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfPMA/pma.cfm.
14. https://www.fda.gov/medicaldevices/deviceregulationandguidance/howtomarketyourdevice/premarketsubmissions/premarketnotification510k/default.htm.
15. https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/pmn.cfm.
16. https://www.fda.gov/medicaldevices/deviceregulationandguidance/howtomarketyourdevice/premarketsubmissions/premarketnotification510k/ucm134575.htm.
17. http://www.gmppublications.com/Part820.htm.
18. https://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/HowtoMarketYourDevice/InvestigationalDeviceExemptionIDE/.
19. https://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/HowtoMarketYourDevice/InvestigationalDeviceExemptionIDE/ucm046164.htm#pre_ide.
20. https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/cfrsearch.cfm?fr=812.2.
21. Kalb S. 2014. https://www.fda.gov/downloads/Training/CDRHLearn/UCM421763.pdf.
22. Phases of medical device IDE trials.
23. https://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/HowtoMarketYourDevice/PremarketSubmissions/HumanitarianDeviceExemption/default.htm.
24. https://www.fda.gov/drugs/developmentapprovalprocess/howdrugsaredevelopedandapproved/approvalapplications/investigationalnewdrugindapplication/default.htm.
25. https://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm.
26. https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?CFRPart=312.
27. https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfCFR/CFRSearch.cfm?CFRPart=312&showFR=1&subpartNode=21:5.0.1.1.3.2.
28. https://www.fda.gov/drugs/developmentapprovalprocess/smallbusinessassistance/ucm069906.htm.
29. https://www.fda.gov/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/InvestigationalNewDrugINDApplication/ucm226358.htm.
30. https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?CFRPart=316.
31. https://www.fda.gov/forindustry/developingproductsforrarediseasesconditions/default.htm.
32. https://www.mastercontrol.com/gxp-lifeline/medical-device-clinical-trials-how-do-they-compare-with-drug-trials-/.

Chapter 9
Statistics: Setting the Stage

M. Abdullah Arain, Adil H. Haider, and Zain G. Hashmi

9.1 Introduction

This chapter introduces fundamental statistical concepts in the design of clinical trials that must be considered early on in the planning phases of a project.

A clinical trial is a prospective experimental research study where participants are assigned to one or more intervention arms to assess the impact of those interventions on various predefined health outcomes [1]. Clinical trials are often conducted to definitively answer clinically relevant questions while overcoming many of the limitations of observational studies. In the hierarchy of evidence-based medicine (EBM), results from clinical trials are considered the highest-quality evidence to drive clinical practice [2]. However, shortfalls in study design, methodology, data analysis, and result interpretation can potentially undermine the value of these studies and reduce their applicability to the general practice. This chapter introduces fundamental statistical concepts in the design of clinical trials that must be considered early on in the planning phases of a project. Over the next few pages, we will dive into the concepts of study design in surgical trials, randomization, allocation sequences, allocation concealment, and blinding and will also touch upon non-randomized trials, pragmatic trials, and superiority and inferiority trials. Eventually, statistical errors and power of the study will be discussed along with sample size considerations.

M. A. Arain · A. H. Haider (✉)
Aga Khan University Medical College, Karachi, Pakistan
e-mail: adil.haider@aku.edu

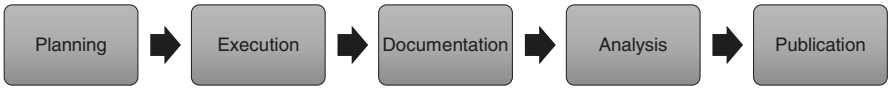
Z. G. Hashmi
Department of Surgery, Sinai Hospital of Baltimore, Baltimore, MD, USA
e-mail: mhashmi@lifebridgehealth.org

T. M. Pawlik, J. A. Sosa (eds.), *Clinical Trials*, Success in Academic Surgery,
https://doi.org/10.1007/978-3-030-35488-6_9

9.2 Setting the Stage

The two important considerations in the development of a clinical trial include (1) asking a clinically important question and (2) having a statistically sound study design to help answer that question. While statistical nuances must be appreciated and understood at each stage of the trial, ignoring them at the design stage can have disastrous downstream consequences. In this regard, only a carefully designed trial that advances clinical knowledge can justify its high costs and resource-intensiveness.

A clinical trial can generally be divided into five phases: planning, execution, documentation, analysis, and publication. The term study design is often used in medical literature to assign an appropriate type to the study while it in fact should refer to the overall planning of all the five phases that together yield a comprehensive publication [3]. Since conducting a trial is a resource-intensive and time-consuming affair, it is extremely disappointing to report less than ideal results just because enough attention was not paid while planning the study.



The Consolidated Standards of Reporting Trials (CONSORT) statement is a set of guidelines that was first introduced in 1996, with the goal of standardizing trial design and conduct. This includes a checklist and a flowchart aimed at standardizing the reporting of clinical trials and guiding the authors on how to be clear, complete, and transparent about their reporting. The statement has undergone several improvements in the last decade to not only aid the reporting of information but also help the readers, reviewers, and scientific journal editors to understand the trial's design, conduct, statistical analysis, and interpretation and critically appraise the publication. The checklist has become a standard of practice when reporting trial data and has been adopted by more than 500 medical journals. The statement also provides a uniformity to the clinical trial literature for later researchers to conveniently select reliable, relevant, and valid studies to include in systematic reviews and meta-analyses [4].

The statistical aspects of a study design include the type of study, sampling, collection of data, and measurement of outcomes/endpoints [5].

9.3 Study Design

Clinical trials fall under the experimental arm of analytical studies investigating the effect of an intervention on the study population. Randomized controlled trials, when designed, carried out, and reported appropriately, represent the most rigorous method of hypothesis testing and are a gold standard in assessing effectiveness of healthcare interventions. However, the results from these trials can suffer from residual biases, especially if adequate methodological diligence is not ensured [6].

Additionally, even though randomization greatly increases the validity of a study, not all clinical trials can be randomized. This is especially true for surgical trials, since the decision to undergo a surgical procedure, when an indication for that procedure exists, is not something that the patients would be willing to leave up to randomization [7]. Outcomes from inadequately designed trials cannot only misguide physicians making treatment decisions for patients at an individual level but also misinform policy makers devising a national public health policy [4]. To overcome these potential pitfalls, a sound understanding of both randomized and non-randomized trials is necessary.

9.4 Randomization

Randomization is the process by which all the study participants possess an equal chance of being assigned to either the experimental or the control groups. This not only removes any selection biases that can potentially impact the outcomes by randomly distributing the patient characteristics between the groups but also equilibrates all confounding factors yielding a control group that is almost exactly congruent to the treatment group. Any disproportionate assignment to either group would skew the results, by introducing conscious and unconscious prejudices, and make the conclusion invalid. Therefore, any subsequent difference in outcomes between the groups can be demonstrated, after being evaluated by the use of probability theory and the level of significance between the different outcomes, to either be the result of difference in intervention or be merely due to chance alone [2, 7, 8]. Randomization also enables multiple levels of blinding (masking) of the intervention from the stakeholding parties like the researchers, participants, and evaluators [8].

There are many processes that can be used to randomize the participants. The aim of each process is to limit bias and to assemble a similar cohort of individuals between groups, and therefore the process should only be administered to the individuals who agree to participate in the study to ensure the purity of the process.

9.4.1 Fixed Allocation Randomization

Randomization by *fixed allocation* means that the assignments to the separate groups would be made at a predefined probability. Usually this proportion is set at an equal allocation (1:1); however, some situations may allow or even necessitate an unequal allocation (2:1). Some researchers argue that unequal allocation is not consistent with the true equipoise of RCTs and tends to introduce a bias to the results; however, others argue that a 2:1 allocation would have minimal effect on the power of the study but could potentially, with a fixed sample size, decrease the cost of the trial significantly. It can, in fact, increase the power of the study by registering a bigger sample size if the funding of the trial is fixed [7, 9]. For example, the

Scandinavian Simvastatin Study for coronary heart disease prevention, with a fixed sample size and a 2:1 allocation ratio, showed a 3% decrease in power while saving 34% of the cost [10].

9.4.1.1 Simple Randomization

Considered to be the most elementary of the randomization processes, sometimes being as basic as rolling a fair dice or tossing a fair coin, simple randomization conserves the absolute unpredictability and bias prevention of each intervention allotment and thus outclasses all other methods of allocation generation, irrespective of their sophistication and complexity. However, this unpredictability can sometimes become a challenge when the sample size is small because there is then a higher likelihood of disparity in group allocation by chance alone. This disparity diminishes as sample size increases [2, 8].

Fair coin tossing, dice rolling, and shuffled deck of cards dealing are examples of manual methods of drawing lots. These methods, though theoretically ideal for random allocation of intervention, are practically susceptible to non-random contamination. A series of tosses with identical outputs could entice the researchers to intervene with the result of the toss because they perceive the randomness of these results to be non-random. These methods also are difficult to implement and do not leave an audit trail and therefore are not the recommended methods of random sampling. Instead, sequence generation by a table of random numbers or computer generation of random numbers is reliable, unpredictable, reproducible, and easily traceable and should be confidently used in trials [8].

Many investigators have a less than ideal understanding of randomization and frequently assume non-random approaches like alternate assignment and haphazard sampling to be random. Quasi-random is a term commonly used to refer to the assignment of intervention groups based on pre-intervention tests, and while the term may have the word random in it, it serves no more than a misnomer for an approach which completely goes against the ideology of randomization. Assignments based on medical record numbers or date of birth (where odd numbers are placed in one group and even in the other) or alternate assignments (e.g., ABABAB) are non-random methods that are mostly mistaken as random. Systematic sampling should not be considered randomized sampling as well because the outcomes are not, theoretically and practically, based on chance alone. Any study not detailing its randomization process or defining its randomization by a non-random method should be approached with caution [8].

9.4.1.2 Restricted Randomization

Restricted randomization (also called blocked randomization) benefits the researchers who require an equal group size in between groups. It restricts large imbalances in sample size by influencing the acquisition of an allocation sequence that could

lead to unequal group sizes. The most commonly used variation of restricted randomization is by random combinations of equal-sized blocks. Participants are examined in blocks of, for example, four individuals at one time. Using this block size will yield six possible combinations of 2 As and 2 Bs in each block (AABB, BBAA, ABAB, BABA, ABBA, BAAB). A random number sequence will then be utilized to select one out of the six blocks, and the sequence of allocation in that particular block is followed for the first four participants. Subsequently, one of these six block combinations will be randomly selected again and its allocation sequence followed for the upcoming four participants and so forth [2, 8]. The downside to this is that some studies could develop an extremely strict exclusion criteria, to keep the population pool comparable, but would end up markedly regulating the participant enrollment and jeopardizing the generalizability of the results [7]. The random allocation rule is another form of restricted randomization where the eventual group sizes would be equal. Often, after the selection of total sample size, a subset is assigned Group A and the remaining end up being Group B. It can be explained by placing in a bowl 100 balls labeled "A" and 100 balls labeled "B" for a total sample size of 200. Then one ball will be drawn at random without replacement and the participant placed in the corresponding group [8].

9.4.1.3 Stratified Randomization

While striving to remove selection bias, randomization tends to establish unwanted chance imbalances. These imbalances can be prevented by dividing the population into strata of prognostic factors (e.g., smokers and nonsmokers). These stratified groups would then undergo blocking randomization to finally yield separate block randomization sequences for the different combinations of prognostic factors. Stratification without blocking would serve no purpose. Even though it is a valid and useful method to curb chance imbalances, the complex stratification procedure provides little benefit in large-scale studies where the effect of chance selection eventually gets balanced on its own. The complicated process can become overwhelming during participant enrollment and sometimes can be a limiting factor for trial collaborations. A possible benefit of stratified randomization is stratification by center in multicenter trials. For small trials, stratification not only provides proper balance between sample groups but also raises the power and precision of the study [2, 7, 8].

9.4.2 Adaptive Allocation Randomization

Adaptive randomization allows for changes in the probability of allocation to the intervention groups with the passage of time. A problem commonly faced with small-sized trials is that simple randomization (with or without stratification of important prognostic variables) results in imbalance of covariates in the

intervention groups. This can lead to incorrect interpretation of research outcomes [7, 11, 12].

Baseline adaptive randomization, like the covariate adaptive randomization (minimization), though in itself is of a non-random nature, is an acceptable and valid alternative mode of randomization in such scenarios. Here, all important prognostic factors (covariates) are identified before the initiation of the trial, and every new participant is assigned sequentially to the specific intervention group by relating their covariates to the already specified ones while keeping in mind the previous allocations to each group. This method achieves balanced groups with respect to numbers and covariates [7, 11, 12].

Response adaptive randomization allows for adjustments in group allocations by evaluating intervention responses. Of these, the "play the winner" style assigns the succeeding participant by relying on the previous participant's response to the intervention. A positive response places the successive participant in the same group, and a negative response shifts the next assignment to the other group. In the "two-armed bandit," the probability of positive results keeps on adjusting as the outcome for each participant is added to the count and more and more participants are added to the group with the superior intervention [7].

9.5 Allocation Concealment

After randomization of the study population by allocation sequences, the next crucial step is to implement the allocation sequence in an impartial and unbiased fashion by concealing the sequence until the patient assignments to the intervention group. Allocation concealment refers to the prevention of foreknowledge of the treatment assignment, thus shielding those who enroll participants from being influenced by this knowledge [4]. An unconcealed study defeats the entire purpose of randomization. Allocation concealment is a widely misunderstood concept. Many researchers delve into conversations about randomization techniques when discussing it and some consider it to be related to blinding. Both of these concepts are inaccurate. Allocation concealment, in reality, is the approach used to implement the allocation sequences generated by randomization. Individuals admitting participants to the study should enroll individuals without having any knowledge regarding allocation of intervention groups for the participants. Any such knowledge has a potential to introduce bias and exaggerate treatment effects producing greatly heterogeneous results than should be expected [2, 13]. Popular methods to conceal allocation include use of opaque envelopes to assign groups, a method not widely recommended, or the use of distance randomization, where allocation sequences are handled by a central randomization service and the investigators have to contact the service for each enrolled participant, placing a gap between recruiters and group allocators [2].

9.6 Blinding

After randomization, there would be a group receiving the (new) intervention and another being administered the control, which could be the already existing standard of care or a placebo. Clinical trials always pose a risk that the knowledge of the benefit of the intervention in the stakeholders (e.g., participants, investigators, analysts) can definitely introduce biases into the study and greatly impact the trial outcomes leading to unacceptable results. The participants, if informed of their assigned intervention, with a previous understanding of the advantages of that intervention, could be led to report positive results even if they did not feel a difference. The investigator, if informed of the intervention group, with a previous perception of potential drawbacks of the intervention, would be led to record negative results even if none actually existed. When analysts have knowledge about which groups they are analyzing, a previous inclination to either one could potentially lead them to over-analyze or under-analyze to produce results that agree with their inclination. Knowledge of the groups can potentially change the delivery of care to either group to adjust for a conceived limitation that the group suffers.

To curtail this probable bias, the trial can be blinded; that is, the participants, investigators, and data handlers can be prevented from knowing which participant belongs to which group, thus preventing the stakeholders from projecting their expected outcomes onto the actual results.

A trial may be single-blinded, where the participants in the group do not know details of their assigned intervention; double-blinded, where both the participants and investigators are kept unaware of the assigned intervention; or triple-blinded, where in addition to the participants and investigators, the data analysts are also kept ignorant of the assignments. Rapid un-blinding should be possible in the design of the study to counter any major harmful effects [2, 14, 15].

It is important to distinguish between allocation concealment, which happens before the randomization process is begun to nullify selection bias, and blinding, which happens post-randomization and reduces detection and performance bias. The overall bias reduction is more significant with allocation concealment than with blinding, and the best approach is to employ both when conducting a study [2, 16].

9.7 Non-randomized Studies

In non-randomly assigned controlled studies, two groups are analyzed against each other, one receiving the new intervention while other being the control. This is very similar to an RCT, except the groups are assigned in a non-random manner. Instances where such a study is acceptable could be when the practicality of large-scale administration of the study is a likely impediment or when the logistic requisites of a standard RCT cannot be fulfilled and concerns of costing and patient acceptability

become the possible limiting factors. However, the study groups being evaluated still need to be comparable.

Non-randomized studies are discouraged in circumstances where numerous confounding variables are being assessed, the endpoints are multifactorial, or there is a lack of evidence in terms of what outcomes to expect. If the researcher is able to identify the confounders and adjust the analysis accordingly, the evidence produced by these studies would be acceptable given the constraints of performing an RCT in a similar situation.

9.7.1 Advantages of the Non-randomized Trial

The benefit of having a control group even in a non-randomized study can never be underestimated. The control group helps maintain the internal validity of the study by nullifying the impact of temporal trends (other aspects of disease and its care), the regression to the mean (outlying values moderating over time), and the learning curve of the surgeon on the outcomes of the study and facilitates the investigator to deal with these elements during the design of the study. The use of non-randomized control can also increase the generalizability of the study by enrolling a heterogeneous population spread across multiple providers.

9.7.2 Disadvantages of the Non-randomized Trial

The most elementary flaw of non-randomized trials is the confounding bias. The direction of this bias is quite variable and unpredictable. Bias introduced by just selectively registering healthier or sicker patients can turn the results in favor or against the intervention, respectively. Therefore, any "hand-picking" of the subjects must be avoided if possible. These studies also fail to account for social, cultural, economic, and clinical variables which have a potential to affect the outcomes. For a better internal and external validity of the study, it must be replicable and be adjustable to various clinical settings for it to have the desired impact [17].

9.7.3 Examples of Non-randomized Trials

Historical control studies trials are an example of non-randomized trials where the intervention group is compared to a previously assessed historical control group. In this method, everyone receives the intervention. However, they may be limited by changes in diagnostic/therapeutic approaches that accrue over time and thus inherent biases can arise. An example would be the difficulty ascribing a mortality

difference to an intervention among patients with coronary artery disease versus historical controls.

Withdrawal studies involve placing participants off treatment to assess the actual benefit of a treatment which has never been proven to be of benefit but is somehow common practice. However, this only allows the most stable patients on the treatment to be selected for the study [7].

Concurrent trials include the crossover design where subjects serve as their own internal control. All subjects are used twice, once in the intervention arm and once in the control arm. Randomization for treatment sequence is also carried out. The major advantage of this method is to account for paired comparisons and mitigate variability secondary to inter-individual differences. Carryover effects are an important consideration for crossover studies. These are effects that "carry over" to the upcoming intervention of the trial from the previous intervention. To counter this effect, studies ensure "wash-in" and "wash out" periods for the succeeding and preceding interventions, respectively. Usually, if more than one intervention is being compared, then a Latin square matrix ($n \times n$) is utilized to ensure that every succeeding intervention is preceded or followed by any other intervention just once. It is believed that this fixed concatenation provides a better control over the carryover effects than by randomization [14].

Factorial study designs commonly involve two interventions to be assessed against the control and a 2 × 2 design is a commonly used factorial design. An important assumption that these studies make is that interventions X and Y independently have no interactions with each other [7].

Control	X + Y
X + Control	Y + Control

9.8 Special Considerations for Surgical Trials

Surgical trials can be classified as those evaluating minor changes in surgical technique, major changes in surgical technique, or surgical versus non-surgical treatments [7]. A common source of error in all these situations can be attributed to the inherent technical variability in the performance of procedures; attempts must be made to mitigate and account for these effects. This especially holds true for novel procedures where the trialist must account for the "learning curve," which entails to the experience of the surgeon on the new procedure which would impact patient selection, operative skills, post-operative care, and additional medical therapy. One way to prevent this is to postpone the initiation of such a trial until adequate expertise has been achieved. Failure to do so may result in an elevated risk of adverse effects and could potentially bias the final results against the new intervention [18].

The underlying principle of equipoise that there must exist a true uncertainty about an intervention's effect in order to justify a clinical trial is even more important for surgical trials. Many patients are concerned about enrolling in surgical trials, especially those involving a novel surgical procedure. A careful explanation of the trial's intents and purpose rooted in equipoise often helps allay these concerns and should always be employed at the time of consent.

Additionally, it may not always be practical to maintain blinding in surgical trials. For example, a comparison of open versus minimally invasive techniques may be hard to blind. However, minor variations in surgical techniques may be amenable to various blinding methods. Moreover, keeping the analytical teams remote from clinical interaction may help maintain blinding in certain situations. Irrespective, careful attention must be paid to enforce and maintain blinding whenever possible.

9.9 Pragmatic Trials and Comparative Effectiveness Research

Randomized controlled trials are conducted to assess if intervention has a biologic impact under strict controlled settings. These trials aim to demonstrate the "potential" of a treatment. Pragmatic controlled trials (PCTs), also called effectiveness trials, are conducted to measure the effectiveness of a treatment in a real-world setting. These trials aim to reach a maximum generalizability of the results while also making sure that the differences in outcomes are a result of the intervention and not due to chance or confounders. In other words, PCTs strive to maintain high external and internal validity [16].

A proposed distinction between explanatory trials (RCTs) and pragmatic trials is that RCTs confirm a physiological or clinical hypothesis while pragmatic trials guide a clinical or policy decision by providing evidence for adoption of the intervention into real-world clinical practice. The need to distinguish between the two became necessary when it was realized that many trials did not adequately inform practice because they were optimized to determine efficacy rather than effectiveness. Since RCTs are performed with relatively small sample sizes, at locations where experienced investigators are conducting these studies with a highly selected population of participants, they could potentially be overestimating benefits and underestimating harm.

The Pragmatic–Explanatory Continuum Indicator Summary (PRECIS) tool attempts to clarify the concept of pragmatism and provides a guide, scoring system, and graphical representation of the pragmatic features of a trial. Included variables are the recruitment of investigators and participants, the intervention and its delivery, follow-up, and the nature, determination, and analysis of outcomes. Most trials could be deemed pragmatic with regard to at least one of these dimensions, but very few end up being pragmatic in all areas [19].

In pragmatic trials, few restrictions are placed on the inclusion criteria to reflect the variation that would exist in the general patient population to ensure generalizability. A larger sample size is needed to cater to the heterogeneity of the population characteristics. These trials usually compare a new treatment to an already existing standard of care rather than a placebo and therefore are preferable as surgical trials where using a placebo or a sham intervention could be considered unethical. Pragmatic trials allow surgeons the flexibility (within set constraints of a real-world setting) to employ their own approaches to the various patients, while also implementing the intervention under trial to the randomly assigned patient. This flexibility in the pragmatic protocols enables academic surgeons to have their conventional practice while also doing research within a defined framework. Instead of measuring surrogate and objective outcomes, like in RCTs, the pragmatic trials focus on patient-centric outcomes like improvement in quality of life (QoL) and follow-up of patients for a longer duration of time. Surgical trials tend to have features of both explanatory trials and pragmatic trials and therefore exist along the continuum of the two designs. Trials designed to eventually aid the clinician to make the best possible decisions for their patients will prove to be most useful [16].

Comparative effectiveness research (CER) differs from clinical trials (especially pragmatic clinical trials) in that it is the conduct and generation of evidence, which incorporates results from observational and experimental researches, including RCTs, to compare the benefits and harms of different interventions to prevent, diagnose, treat, and monitor a clinical condition in the everyday settings and helps improve the delivery of care. A clinical trial is not CER in and of itself; however, CER uses results from clinical trials to inform clinical care. In general, CER aims to assist consumers, clinicians, purchasers, and policy makers to make informed decisions and improve healthcare outcomes for individual patients and patient populations [20].

9.10 Superiority/Inferiority Trials

The type of RCT being conducted depends on the aim of the trial itself. If the aim is demonstrating that the intervention (E) is superior to the control (C), then it is considered to be a superiority trial and the statistical tests executed are the superiority tests. If the results are significant, it could be concluded that intervention produces significantly better outcomes than the control. On the other hand, non-significant results are difficult to categorize as they obviously do not show superiority but also essentially do not show that the intervention was not as equally effective as the control. In fact, there will always exist a small difference in effects when two treatments are non-identical and yet the primary effect could very much be similar. The different interventions could also have the same primary effect but have secondary qualities that could make one more preferable over the other, and it was situations like this that have led to the inception of non-inferiority trials (NITs). Non-inferiority trials could be performed if it is ethically not appropriate to create a placebo group,

which is a problem that is commonly encountered in surgical trials. NITs could also be a viable option if the primary outcomes between the groups are expected to be similar but secondary outcomes or safety profiles are anticipated to be better with the new intervention or the new treatment is cheaper and/or easier to administer and is more likely to be applicable in real-life situations [21].

9.11 Outcomes

All RCTs assess response variables or endpoints (outcomes) for which the groups are analyzed against each other. RCTs generally have a diverse set of variables being assessed and the investigator must define and specify the importance of each output variable during the design phase of the study. Each variable could be one of three types of response variables that are measured: dichotomous (measuring event rates), continuous (measuring mean values), and time to event (measuring hazard rates) [4, 7].

The primary endpoint is the predefined response variable that holds the greatest significance for all the involved parties (the patients, investigators, financers, and policy makers) and is mostly the treatment effect variable used when calculating the sample size. It is likely that a trial could be assessing multiple primary variables; however, this comes with its own issues of result interpretation and is discouraged. All other outcomes being evaluated are termed secondary outcomes, and these could be outcomes that were expected and observed and also those which were not expected but were still observed. Adverse effects should always be given importance, irrespective of their status as primary or secondary outcomes. The variable should be defined in way that a third party reading the study should be able to understand and use the same variables. Appropriate use of previously validated guidelines or scales is recommended to enhance the quality of the measurement and make future comparisons possible.

Any unplanned digression from the initially approved protocol must be reported. All changes to the selection criteria, intervention itself, data recording, analytical adjustments, and the reported outcomes ought to be clearly reported [4].

9.12 Types of Errors and Statistical Power

Statistical considerations must be made early during the planning phases of a clinical trial in order to truly understand the results and avoid pitfalls in sample size calculations, power, and statistical significance. Usually the null hypothesis, which states that no observed difference exists between two (or more) groups, is tested using appropriate statistical tests. Several types of errors can occur when interpreting these results which are discussed here.

9.12.1 Type I Error (α)

This is the probability of detecting a statistically significant difference when in fact no difference exists, that is, the chance of a false-positive result.

9.12.2 P-Value (p) and Significance Level (Alpha)

P-value is the probability of a type I error, that is, the probability of detecting a difference as large (or larger) as the actual difference observed given that the null hypothesis is true. The significance level (alpha) refers to the probability that is decided a priori, while *p*-value refers to calculated value obtained after performing a statistical test. Typically, the null hypothesis is rejected if the *p*-value is less than the chosen alpha. Alpha is often chosen arbitrarily but is conventionally set at 0.05 (1 in 20 chance of being incorrect) or 0.01 (1 in 100 chance of being incorrect). In general, the larger the alpha, the larger the required sample size.

9.12.3 Type II Error (β)

This is the probability of not detecting a statistically significant difference when in fact a difference truly exists, that is, the chance of a false-negative result.

9.12.4 Power (1–β)

Power is the probability of detecting a statistically significant difference when in fact a difference truly exists, or alternatively, the probability of rejecting a null hypothesis when it is false. In simpler terms, power quantifies the ability of a study to find true differences. Beta depends on alpha, the sample size, and the measure of true difference between variables (delta). In general, the higher the power, the larger the required sample size. Usually, alpha is set at 0.05 or 0.01 and beta is set at 0.90 or 0.95, while delta and sample size are variable. Delta is typically based on prior research findings and is set at the minimal level at which the differences between groups still remain clinically meaningful.

	When null hypothesis (H_0) is	
Statistical result	True	False
Reject null hypothesis (H_0)	Type I error (α)	Power (1-β), correct result
Fail to reject null hypothesis	Correct result	Type II error (β)

9.13 Sample Size Considerations

A clinical trial should have a sufficient sample size that would ensure the detection of a statistically significant, clinically meaningful effect of an intervention if in fact it truly exists. The end goal is to determine the most conservative sample size in order to avoid overestimates (failure to enroll, high costs) and underestimates (inconclusive results).

The exact details of accurate sample size calculation are beyond the scope of this discussion. However, a few broad concepts need to be understood in order to have a clear view of its mechanics. As alluded to before, in calculating sample sizes, beta and alpha are usually set by convention, while delta is estimated based on prior research. The larger the delta, the smaller the sample size needed to detect a true difference. These differences are typically tested using two-sided tests to detect differences in either direction (since a new treatment may perform better or worse than standard of care/placebo). The significance level used for sample size calculations for two-sided tests is twice that of a one-sided test; therefore, the choice of hypothesis testing has a bearing on sample size calculation. Another design consideration of sample size calculations for clinical trials is the allocation ratio for the probability of assignment to the treatment groups. Most researchers choose a 1:1 allocation ratio for their trials, which means the probability of assignment to the two groups is equal. Although a 1:1 allocation ratio usually maximizes trial power, ratios up to 2:1 minimally reduce the power [4, 22] and require a higher sample size. Lastly, many clinical trials routinely call for interim analyses to serve as an early warning system. If the treatment is overwhelmingly useful or harmful or an expected difference does not result, the trial may need to be stopped earlier due to safety concerns or to preserve resources. When performing these interim analyses, careful consideration of the sample size and initial significance level must be undertaken and adjusted for since the rate of incorrectly rejecting the null hypothesis will be larger.

9.14 Conclusion

Clinical trials, when planned and executed correctly, represent one of the best strategies to determine the clinical benefit or harm as a direct consequence of a new therapeutic intervention. Meticulous, upfront attention during the early design phases of a clinical trial will not only save precious resources, but also will result in the advancement of clinical therapeutics.

References

1. Winter K, Pugh SL, editors. An investigator's introduction to statistical considerations in clinical trials. Urologic oncology: seminars and original investigations. Amsterdam: Elsevier; 2019.
2. Akobeng AK. Understanding randomised controlled trials. Arch Dis Childhood. 2005;90(8):840–4.

3. Rohrig B, du Prel JB, Blettner M. Study design in medical research: part 2 of a series on the evaluation of scientific publications. Dtsch Arztebl Int. 2009;106(11):184–9.
4. Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. Int J Surg. 2012;10(1):28–55.
5. Van Belle G, Fisher LB. Biostatistics: a methodology for the health sciences. 2nd ed. Hoboken: Wiley; 2004. p. 15–9.
6. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ. 2010;340:c332.
7. Pawlik TM, Sosa JA. Success in academic surgery: clinical trials. 1st Edition. Springer. 2014. Pg 14, section 2.3; ISBN 978-1-4471-4679-7
8. Schulz KF, Grimes DA. Generation of allocation sequences in randomised trials: chance, not choice. Lancet. 2002;359(9305):515–9.
9. Peckham E, Brabyn S, Cook L, Devlin T, Dumville J, Torgerson DJ. The use of unequal randomisation in clinical trials—an update. Contemp Clin Trials. 2015;45:113–22.
10. Torgerson DJ, Campbell MK. Use of unequal randomisation to aid the economic efficiency of clinical trials. BMJ. 2000;321(7263):759.
11. Suresh K. An overview of randomization techniques: an unbiased assessment of outcome in clinical research. J Hum Reprod Sci. 2011;4(1):8.
12. Scott NW, McPherson GC, Ramsay CR, Campbell MK. The method of minimization for allocation to clinical trials: a review. Control Clin Trial. 2002;23(6):662–74.
13. Schulz KF, Grimes DA. Allocation concealment in randomised trials: defending against deciphering. Lancet. 2002;359(9306):614–8.
14. Spieth PM, Kubasch AS, Penzlin AI, Illigens BM-W, Barlinn K, Siepmann T. Randomized controlled trials—a matter of design. Neuropsy Dis Treat. 2016;12:1341.
15. Kendall JM. Designing a research project: randomised controlled trials and their principles. Emerg Med J. 2003;20(2):164–8.
16. Farrokhyar F, Karanicolas PJ, Thoma A, Simunovic M, Bhandari M, Devereaux P, et al. Randomized controlled trials of surgical interventions. Annal Surg. 2010;251(3):409–16.
17. Axelrod DA, Hayward R. Nonrandomized interventional study designs (quasi-experimental designs). Clinical research methods for surgeons. New York: Springer; 2006. p. 63–76.
18. Thoma A, Farrokhyar F, Bhandari M, Tandan V. Users' guide to the surgical literature. How to assess a randomized controlled trial in surgery. Can J Surg. 2004;47(3):200–8.
19. Ford I, Norrie J. Pragmatic trials. N Engl J Med. 2016;375(5):454–63.
20. Chang TI, Winkelmayer WC. Comparative effectiveness research: what is it and why do we need it in nephrology? Nephrol Dia Transplant. 2012;27(6):2156–61.
21. Lesaffre E. Superiority, equivalence, and non-inferiority trials. Bull NYU Hosp Jt Dis. 2008;66(2)
22. Schulz KF, Grimes DA. Sample size calculations in randomised trials: mandatory and mystical. Lancet. 2005;365(9467):1348–53.

Chapter 10
Clinical Trials: Handling the Data

Douglas S. Swords and Benjamin S. Brooke

10.1 Introduction

Clinical trials constitute a central role in answering questions about the efficacy of different surgical interventions, including new treatments and known interventions that warrant further study or comparison. Randomization is necessary to discern treatment effects for surgical interventions with modest to marginal effect sizes. This process involves prospectively assigning human subjects to one or more interventions in a randomized fashion under conditions that are controlled by the investigator and then evaluating their effects on health-related, behavioral, or patient-reported outcomes. To carry out and interpret the findings from clinical trials, surgical investigators need to understand all the nuances of methodology including how to handle the data and appreciate issues that might influence the analysis. The purpose of this chapter is to review the methodological considerations that are intrinsic to designing, interpreting, and reporting clinical trials.

10.2 Hypothesis Testing

Clinical trials are undertaken to provide data that will help answer scientific questions where the truth is unknown. Yet before undertaking a trial to answer this question, an investigator must first state their scientific hypothesis where they think the truth lies based on existing evidence. Hypothesis testing is an approach for choosing between two competing possibilities and is central to statistical inference.

D. S. Swords · B. S. Brooke (✉)
Utah Interventional Quality and Implementation Research (U-INQUIRE) Group, Department of Surgery, University of Utah School of Medicine, Salt Lake City, UT, USA
e-mail: Douglas.Swords@hsc.utah.edu; Benjamin.Brooke@hsc.utah.edu

T. M. Pawlik, J. A. Sosa (eds.), *Clinical Trials*, Success in Academic Surgery,
https://doi.org/10.1007/978-3-030-35488-6_10

Importantly, it provides an objective framework for making decisions using probabilistic methods rather than simply relying upon subjective impressions.

In clinical trials that compare outcomes between two or more groups, a hypothesis is proposed for the statistical relationship between different groups. The assumption that there is no difference in outcomes between groups is called the null hypothesis (H_0). In comparison, the alternative hypothesis (H_1) is the investigator's scientific hypothesis that specifies a difference between groups and which data is being collected to answer. To test if H_1 is true, an investigator needs to show that the probability that the observed data satisfied the null hypothesis is very unlikely. But erroneous conclusions with regard to the null hypothesis can sometimes occur by chance alone and be categorized into two types of random error: type I and type II.

10.2.1 Type I Errors

Type I errors, also called *alpha errors*, occur when researchers erroneously reject the null hypothesis. Specifically, it is inferred that there is a difference in outcomes when in fact there is no difference between groups. This is considered a false-positive result. Statistical testing is used to quantify the likelihood of a type I error occurring within predetermined probabilities. A *p-value* indicates the probability that observed differences between groups might be due to chance alone. In other words, the difference may not be based on the effect of the intervention being tested. The threshold for statistical significance is conventionally set at a *p*-value of 0.05, signifying that the likelihood of the observed differences being due to chance alone might occur 5 times out of 100 tests. Although a likelihood of 5% falls short of absolute certainty, this level of confidence is generally accepted as scientific proof and used throughout the scientific literature.

Type I errors can occur when the research question and analysis have not been specified a priori, or when multiple statistical tests are performed in a study with several subgroups. A study with over 20 comparisons, for instance, will be expected by chance alone to have at least one false-positive finding with a *p*-value set at 0.05. When 20 or more comparisons are necessary in a given study, a Bonferroni correction or Hochberg sequential procedure can be used to protect against a type I error occurring. The Bonferroni correction works by testing each hypothesis at a significance level that is determined by the alpha level (i.e., *p*-value) divided by the number of comparisons or hypotheses. For example, if a clinical trial was planned to test 20 different hypotheses at a *p*-value of 0.05, the Bonferroni correction would test each comparison at 0.05/20 = 0.0025 to meet statistical significance.

Beyond multiple comparisons in a clinical trial, however, there has been a lot of debate around using the *p*-value threshold of 0.05 for determining statistical significance and type I errors. A recent letter to the journal *Nature* signed by over 800 scientists argues that it's inappropriate to conclude that there is "no difference" or "no association" in a clinical trial just because a *p*-value is larger than a threshold such as 0.05 or, equivalently, because a confidence interval includes zero [1].

These authors' argument is that it is illogical to dichotomize the concept of statistical significance when it is, in reality, a continuum. Dr. John Ioannidis, who has written extensively on the replication crisis in biomedicine, argues that such an approach promotes bias and allows pharmaceutical companies to use the weak suggestion of benefit to promote their products [2]. He recommends lowering standard p-value thresholds from 0.05 to 0.005 [3], which would move about 1/3 of statistically significant results in the biomedical literature from "significant" to "suggestive" [4]. While this debate is unlikely to be definitively solved, it is important that surgeons involved in clinical trials be aware of the merits of both viewpoints.

10.2.2 Type II Errors

Type II errors, also called *beta errors*, occur when researchers erroneously confirm the null hypothesis. In this instance, it is inferred that there is no difference in outcomes when, in reality, a difference exists. This is considered a false-negative result and often arises when the sample size is simply insufficient to detect small but clinically important differences in outcomes. When a study's sample size is too small to detect differences in outcomes between comparison groups, it is said to lack sufficient statistical power. But once a study is complete, no amount of analysis can correct for insufficient statistical power. Before starting a prospective study, researchers should perform a "power calculation," which involves determining the minimum size of a meaningful difference in outcomes and then calculating the number of observations required to show that difference statistically. The sample size of most clinical trials is calculated with a power of 80–90%, meaning that there is a 10–20% chance of not finding a difference when one exists (i.e., *β error)*. Surgeons should be particularly cautious when evaluating studies with null findings, particularly when no power calculation is explicitly reported.

10.2.3 Confidence Intervals

An alternative expression of statistical likelihood is confidence intervals. A confidence interval is a range of values that an investigator can be certain contains the true mean of the population. Confidence intervals can also be defined as showing the range of the observed difference that would be expected if the same study were repeated an infinite number of times. For example, a 95% confidence interval would include the observed difference 95% of the time that the study was repeated. Factors affecting the width of the confidence interval include the size of the sample, the confidence level, and the variability in the sample. When all other factors are equal, a large sample size will tend to produce a better estimate of the population parameter.

10.3 Bias and Error

Assessing the internal validity of a clinical trial requires an understanding of the potential influence of bias and random error on the study results. Bias refers to systematic errors in how study subjects were selected or assessed that result in an inaccurate estimate of the differences in outcomes between comparison groups. In comparison, random error refers to the unpredictable randomness of events that might mislead how the study data is analyzed. The potential adverse effect of bias and error in a clinical trial is that investigators will come to the wrong conclusions about either the beneficial or harmful effects of a given intervention.

10.3.1 Sources of Bias

There are two major categories of bias that can impact a clinical trial: selection bias and observer bias. Either of these biases can arise from attitudes and beliefs among the investigators that may affect how the study is designed or analyzed. As such, investigators should apply methods and techniques that control for these types of biases whenever possible.

Selection bias refers to any imperfection in the process by which subjects are selected for a study. Depending on the specific inclusion and exclusion criteria, a study cohort might have subjects who are not typical of the target population. This can be a limiting factor in the external validity or generalizability of study findings. Selection bias in clinical trials may also result in cohorts that are more or less likely to follow-up or have the outcome of interest. For example, follow-up rates among patients participating in a clinical trial comparing surgery versus medical treatment for a specific condition might be different depending on characteristics of patients within each treatment arm. Selection bias will result if study follow-up is curtailed or dropout rates increased in one group for reasons that are connected to the primary outcomes.

For clinical trials that use restricted randomization techniques such as blocking or stratification by site of recruitment, selection bias can also be a problem if investigators can guess the next allocation with greater than 50% probability. For example, when the randomization procedure is restricted to ensure that an equal number of patients are allocated at each trial site, the probability of the next allocation will depend on the previous allocation. In this case the investigator might be able to guess with a high probability what treatment group the next enrolled patient will receive. In comparison, investigators won't be able to calculate the probability of the next treatment group allocation when randomization is not stratified by site.

Observer bias, also referred to as detection bias or ascertainment bias, is a type of measurement bias that refers to problems caused by the way information about outcomes or other pertinent data is obtained during a clinical trial. Sources of measurement bias may be subtler than selection bias. For example, asking surgeons to

assess surgical site infection (SSI) outcomes in their own patients might result in erroneous reported rates as compared to SSI rates reported by hospital epidemiologists or persons not involved in the trial. Blinding is one of the most common efforts to control measurement bias, where neither the subjects nor assessors are aware what intervention was performed. Clinical trials that used subjective measurement scales such as injury severity scores are also susceptible to observer bias. In this case, a method to control for observer bias is to have objective and standardized study outcome measures.

10.3.2 Random Error

Random error is another form of error in measurement caused by factors which vary from one measurement to another. It is also known as variability (i.e., random variation) or the degree of "noise in the study population." In clinical trials, heterogeneity in the population of study subjects can lead to relatively large random variation or imprecision in study results that are scattered around the mean values. Investigators should be considerate of these elements of random error when analyzing their data.

In clinical trials, random error can usually be accounted for by averaging outcomes over a large number of observations. In other words, it is ensuring that the sample size of comparison groups is adequate. Because random error has no preferred direction, it will ultimately yield a net effect of zero.

10.4 Important Elements of Clinical Trial Design

10.4.1 Random Allocation

Each of the participants in a clinical trial should have an equal chance to be allocated to the treatment interventions or control group. The easiest way to achieve this is through random allocation to one of the study interventions, equivalent to tossing a coin for each patient assignment. In clinical trials, the randomization process usually consists of two steps: (1) generating an unpredictable random sequence for treatment allocation and (2) concealing the sequence in such a way that patients (or investigators) don't know the allocation until they have been formally assigned to a treatment arm of the clinical trial.

There are several different methods used to generate a random sequence for treatment allocation. Simple random allocation may include using a random-numbers table or computer software program that generates a random sequence. However, these methods can lead to unequal group sizes by chance alone, particularly when sample sizes are low. To ensure that treatment group sizes are balanced and statistical power is maximized, procedures can be used such as stratified randomization or permuted-block randomization.

10.4.2 Stratified Randomization

Randomization tends to produce groups which are, on average, similar in distribution of baseline characteristics. However, it is possible through random chance that important prognostic variables may be unequally distributed, particularly in smaller trials. Covariate imbalance due to random chance is therefore an important issue in RCTs, particularly in trials with a small sample size. Covariate adjustment and stratification are two techniques that can be used "after the fact" in the analysis phase to deal with imbalances in prognostic factors. Alternatively, stratified randomization can be used to prevent this issue.

Stratified randomization requires that prognostic factors which an investigator wishes to stratify by be measured at or before randomization. The total number of strata is the product of the number of subgroups (values) for each stratification variable. For example, if one wished to stratify by sex and age (defined as 20–30, 31–50, and 51–70), there would be 6 strata. Stratified randomization involves assigning each participant to a stratum before randomization and then randomizing within that stratum.

10.4.3 Blocked Randomization

Permuted-block randomization, also known as blocked randomization, is another one of the common techniques for balancing patient allocation in a large clinical trial. In this technique, each "block" has a specified number of randomly ordered treatment assignments, and patients are randomized in sequential blocks. For example, in a clinical trial comparing treatment X versus treatment Y where the block size is four, there are six possible ways to make treatment assignments within a block: XXYY, YYXX, XYXY, YXYX, XYYX, and YXXY. The principle of blocking is used to increase the power of treatment comparison by dividing the experimental units into blocks as well as pooling the group differences over blocks. This becomes even more important when the characteristics of patients enrolled into a clinical trial change over time.

10.4.4 Allocation Concealment

Allocation concealment is another key component of the randomization process in a clinical trial. This means that neither investigators nor participants are aware of whether the next eligible participant will be receiving the treatment or control intervention. This should be masked until the time when participants are ready to receive the intervention in order to prevent selection bias. For example, an investigator might decide not to enroll a patient in a study if they know that they are allocated to a control group and have a bias toward the treatment being studied. This situation becomes very important when blinding of interventions is not possible.

10.4.5 Blinding

The purpose of blinding in RCTs is to reduce the potential for selection bias. Selection or information bias may be introduced if either the investigators or participants are aware of who is getting the interventions and who is not. Ideally, clinical trials would always have a double-blind design. However, this is often not possible, especially when surgical interventions are tested. Nevertheless, it is important to understand the strengths and limitations of different types of blinding.

In an un-blinded (i.e., open) trial, both the participant and the investigators know which arm the participant is assigned to. While this design has the advantage of being easier to conduct than other study types, it is susceptible to several sources of bias. For one, participants assigned to the control arm might drop out at higher rates than those assigned to the intervention arm if they (or the investigators) have preconceived notions about the benefits of each treatment arm. When the outcome is a subjective measure, this can result in knowledge of one's treatment assignment influencing the outcome measure. Additionally, un-blinded trials are susceptible to study arm differences in compensatory treatment. That is, investigators may prescribe control arm participants additional treatment to "compensate" for not receiving the intervention of interest.

In a single-blind study, the participants are unaware of which intervention they are receiving, but the investigators *are* aware. Like un-blinded studies, the main advantage of this study design is its simplicity in comparison to double-blind studies. Additionally, the issues of unequal participant dropout due to knowledge of study arm assignment and biased participant reporting of subjective outcomes are minimized. However, an investigator's inherent bias can still affect data collection and assessment of outcomes. Finally, single-blind studies are also susceptible to compensatory treatment, as described above in the section on un-blinded trials.

Double-blind studies are considered the gold standard for testing treatment effects in RCTs. In a double-blind study, neither the participants nor the team of investigators know the intervention assignment until the trial is over. The main advantage of this design is that the risk of bias is minimized because preconceived ideas of both the investigators and participants are minimized.

Triple-blind studies are an extension of the double-blind design. In this design, the committee responsible for monitoring outcomes is additionally blinded to group assignment. This added feature is based on the theoretical advantage that the monitoring committee may be able to adjudicate outcomes more objectively if they are unaware of treatment group assignment. However, such a design can be counterproductive with regard to the monitoring committee's ethical duty to minimize harm to participants. If a triple-design design is employed, the monitoring committee is often given the authority to ask for participant group un-blinding at any time if concerning trends in adverse events develop.

10.4.6 Outcome Ascertainment

The primary and secondary outcomes of clinical trials need to be specified prior to conducting the study. Whenever possible, these predefined outcome measures should be collected by independent observers who are unaware of the allocation and treatment arms to prevent ascertainment bias. This is particularly important when interventions cannot be masked, such as by the presence of surgical incisions or wounds. It is also important that the outcome measures are collected in all randomized patients, with measures taken to minimize missing outcomes as far as possible. A high rate of attrition will lead to reduced confidence in the results and may lead to biased estimates, particularly if attrition is unequally distributed between arms of the clinical trial.

10.5 Statistical Analysis of Outcomes

10.5.1 Choice of Statistical Test

The appropriate statistical test for analysis of the primary outcome and secondary outcomes in clinical trials must be selected according to several factors. These include (1) the number of arms in the trial, (2) the type of outcome data being analyzed, (3) whether the data are correlated, and (4) the number of observations in each comparison group.

The most common type of clinical trial involves two treatment arms and is the focus of this discussion involving statistical tests. However, a few issues regarding the analysis of trials with three or more arms are worth mentioning. First, when interpreting trials with three or more arms, it is important to understand whether "between-group" or "among-group" comparisons are being presented. Among-group comparisons can be difficult to interpret, especially if one is only interested in one comparison, and between-group comparisons are more informative. Second, increasing the number of analyses in a dataset can increase the chance of type I error. In a two-armed trial of A vs. B, there is only one possible comparison. However, in a three-armed trial of A vs. B vs. C, there are three possible comparisons (A vs. B, A vs. C, and B vs. C). As described in Sect. 10.2.1, one approach that can be used to account for the increased chance of type I error is to apply a Bonferroni correction factor. However, this approach is not universal and can increase the chance of a type II error.

The choice of statistical test depends on the data type for the chosen outcome variable. Common types of data and examples of each are shown in Table 10.1. Categorical variables are also known as qualitative variables and can be further categorized as either dichotomous, nominal, or ordinal. Dichotomous variables have only two categories or levels, such as "yes" versus "no" or "male" versus "female." Nominal variables have two or more categories that do not have an intrin-

Table 10.1 Types of outcome variables and examples

Data type	Examples
Dichotomous (aka binary)	1. 30-day readmission 2. In-hospital mortality
Unordered categorical (i.e., nominal)	1. Career outcome from an education intervention (i.e., engineering, medicine, dentistry, law, podiatry) 2. Car brand purchased after seeing an advertisement (i.e., Chevrolet, Ford, Subaru, Jeep)
Ordered categorical (i.e., ordinal)	1. Outcome on a Likert scale 2. Education completed (i.e., some high school, high school graduate, some college, college graduate)
Continuous	1. Blood pressure 2. Weight
Censored time-to-event	1. Overall survival 2. Recurrence-free survival

sic order. In comparison, ordinal variables have two or more categories just like nominal variables, but the categories are ordered or ranked. Finally, continuous variables are also known as quantitative variables and can be further categorized as either interval or ratio scale variables. Intervals are data with equality between one measure and the next measure (i.e., blood pressure), whereas ratio scales are numbers used as measurements that have numeric value (i.e., weight).

A correlated data structure generally refers to one of two situations. First, a clinical trial might wish to measure a repeated measure in the same patient. For example, consider a clinical trial that is designed to measure the weight of participants at baseline and after assignment to an exercise program vs. usual care. It is necessary to account for the fact that participant's weight at follow-up is more similar to their baseline weight than to other subjects' weights (i.e., the data are correlated). Second, patients may be "clustered" within units such as clinics, hospitals, or geographic units. If patients within each unit are more similar to themselves than to the overall sample, then the data structure can also be considered as correlated. This situation is commonly found in cluster randomized clinical trials.

Knowledge of the data type and whether a correlated data structure is present allows selection of the appropriate statistical test, as shown in Table 10.2. A chi-square test is generally preferred for analysis of dichotomous and unordered categorical outcomes so long as the *minimum expected frequency* rule is met. When dealing with a dichotomous outcome, the chi-square test should not be used if $n < 20$. If $20 < n < 40$, the chi-square test should not be used if any expected frequency is less than 5. When $n \geq 40$, three of the expected cell frequencies should be at least 5 and one expected frequency can be as small as 1. When analyzing an unordered categorical outcome with >2 outcome values, the chi-square test can be used if no more than 20% of the cells have expected frequencies <5 and no cell has an expected frequency <1. If the *minimum expected frequency* is not met, then Fisher's exact test and the Fisher-Freeman-Halton test should instead be used for dichotomous and unordered categorical outcomes, respectively.

Table 10.2 Statistical tests that can be used to compare different variables

Data type	Two independent groups	Two correlated samples
Dichotomous (aka binary)	Chi-square test or Fisher's exact test	McNemar test
Unordered categorical (i.e., nominal)	Chi-square test or Fisher-Freeman-Halton test	Stuart-Maxwell test
Ordered categorical (i.e., ordinal)	Wilcoxon-Mann-Whitney test	Wilcoxon sign rank test
Continuous	Independent groups *t*-test	Paired *t*-test
Censored time-to-event	Log-rank test	Shared-frailty Cox regression

10.5.2 Analysis of Time-to-Event Outcomes

Survival analysis methods are essential for analyzing clinical trials in which participants have variable lengths of follow-up. Simple comparison of event rates is inappropriate when the length of observation is different for each participant. Survival curves are generated to permit comparison of the follow-up experience of all participants, accounting for both follow-up times and dropouts.

It is important to understand the notation that is used in generating a survival curve. Whether the outcome event has occurred must be recorded for each participant. Additionally, the amount of follow-up time is recorded for each participant. Since participants are likely to enter the trial at different times, it is usually easiest to record the date of entry and the date of last follow-up. The date of entry can then be subtracted from the date of last follow-up during the analysis phase to generate the follow-up time. Participants who do not experience the event before last follow-up are said to be *censored*.

The Kaplan-Meier estimate (i.e., the product limit estimate) uses the concept of conditional probability estimates to estimate survival curves in data with censored observations. The example in Table 10.3 will be used to illustrate the Kaplan-Meier method. In this example, 50 participants were entered on January 1, 2015, and another 50 were entered on January 1, 2016. Therefore, there were unequal periods of follow-up between the two groups when the data were analyzed on January 1, 2017. The 1-year survival rate, which utilizes the data from both of the groups, is $(40 + 35)/(50 + 50) = 75\%$. The 2-year survival rate uses only the data for the group that enrolled in 2015 (i.e., the patients enrolled in 2016 are ignored). The 2-year survival rate is the product of the 1-year survival rate times the probability of surviving the second year: $75 \times (30/40) = 56\%$.

Trial protocols should specify whether time-to-event outcomes will be compared using total curve comparisons or point-by-point comparisons. The log-rank test is a total curve comparison or a statistical test of the overall survival experience. Conversely, a trial could state that the endpoint is the survival rate at a particular point in time after randomization (i.e., 2-year survival). In general, point-by-point comparisons are not recommended unless there is a particular reason to do so.

Table 10.3 Assessment of follow-up using Kaplan-Meier method

Years of follow-up		Entry date	
		Jan 1, 2015	Jan 1, 2016
1	Participants entered	50	50
	First-year deaths	10	15
	First-year survivors	40	35
2	Participants entered	40	
	First-year deaths	10	
	First-year survivors	30	

Survival curves will frequently come close together at a certain time point even if they are fairly separate over most of the follow-up period (and vice versa). Whole curve comparisons represent all of the available follow-up data.

10.5.3 Relative Effect Measures

Investigators commonly wish to present a relative effect measure in addition to raw outcome data and results of significance testing. Dichotomous outcomes can be expressed using both odds ratios (ORs) and relative risk ratios (RRRs). The odds of an event are the ratio of the probability of an outcome occurring to the probability of it not occurring. If the rate of mortality is 20%, the odds of mortality would be 0.2/(1–0.2) = 0.25. When the probability is small, the odds will be very similar to the probability. For example, for a probability of 0.07, the odds are 0.07/(1 – 0.07) = 0.0753. ORs are the most common relative effect measure in the medical literature. One benefit of the OR is that it is easy to test the statistical strength of an association. Logistic regression tests whether the parameter (log odds) equals 0, which corresponds to whether the OR equals 1.0. ORs are usually reported with 95% confidence intervals (CIs); if the 95% CI does not include 1.0, then the OR is considered significant at $p < 0.05$.

Despite ORs being the most commonly used relative effect measure, they do have limitations [5]. First, although the OR frames events in terms of odds rather than probability, clinicians commonly misinterpret the OR as an RRR. Although ORs approximate the RRR with rare outcomes (those occurring <10% of the time), ORs are always inflated estimates of relative risk with more common outcomes. For example, an event that occurred at a rate of 50% in the treatment arm and 33% in the control arm would have an OR of [0.5/(1 – 0.5)]/[0.33/(1 – 0.33)] = 1/0.493 ~ 2.0. The corresponding RRR would be 0.5/0.33 = 1.52.

A second and less well-understood limitation of using OR relates to the fact that the magnitude of this effect measure is scaled by an arbitrary factor (the square root of the variance of the unexplained part of the binary outcome) [5, 6]. Therefore, adding more explanatory variables to a logistic regression model can artificially

increase the OR for the treatment variable because the scaling factor will be smaller. This curiosity means that it is impossible to compare ORs between models that adjust for different covariates, which presents difficulties when one wishes to conduct a meta-analysis. For these reasons, many have called for authors to report RRRs instead of ORs. This can be done using a number of methods, but we prefer using marginal standardization. In this technique, a logistic regression model is fitted and then a post-estimation command is run to obtain the RRR with 95% CIs. The adjrr package in Stata does this nicely [7], and other statistical packages offer similar suites.

The hazard ratio (HR) is the interval-specific or "instantaneous" risk of the outcome. Cox regression does not assume equal follow-up time for each subject and allows the number at risk (denominator) to decrease across the follow-up time. The hazard ratio is, in fact, a weighted average of the interval-specific risk ratios. The p-value obtained during unadjusted Cox regression is not identical to the log-rank test (although they are usually close in value). Therefore, the p from the log-rank test should still be reported in analysis of time-to-event outcomes even if an HR is additionally calculated.

10.5.4 Absolute Effect Measures

In clinical trials, it is helpful to readers to present absolute effect measures in addition to relative effect measures. Neither the relative measure alone nor the absolute measure alone gives a complete picture of the effect and its implications. Many readers tend to overestimate effect size when only a relative effect measure is reported. Furthermore, the OR and RRR are dependent on the underlying rate, whereas adjusted risk difference (ARR) is not. An unadjusted and adjusted ARD can easily be computed using post-estimation commands after fitting a logistic regression model [7].

There is also a corresponding absolute effect measure for time-to-event data. The difference in restricted mean survival times (RMSTD) compares the mean survival times between groups up to a fixed point [8]. The RMSTD questions how much longer, on average, participants receiving an intervention live over a fixed time horizon. In an RCT, the results are more likely to be correctly interpreted when both the HR and the RMSTD are presented [9].

10.6 Intention-to-Treat Analyses

Intention-to-treat analysis (ITT) is a technique used in RCTs, where patients are compared within the groups to which they were initially randomized to. This is independent of the treatment they actually received, or irrespective of whether they dropped out of the study or violated the study protocol. In other words, it constitutes

an analysis of the results based on the treatment arm to which the patients belong due to the initial random allocation and not on the treatment actually received (active or placebo). ITT analysis permits the pragmatic evaluation of the benefit of a treatment change and not the potential benefit in patients getting the pre-planned allocated treatment only. Full application of the ITT principle is only possible in those circumstances where all results from all patients are available.

For superiority trials, the ITT analysis should be considered the most important analysis. The ITT analysis stands in contrast to a per-protocol analysis, where patients are analyzed by the treatment that they actually received. Aside from drop-out from the study, the main difference between the ITT analysis and the per-protocol analysis is *adherence*. Patients commonly do not actually receive all of a treatment which they were randomized to because their status declines and they are unable to tolerate a full treatment course. Therefore, per-protocol analyses are subject to selection biases that ITT analyses are not. While it is important to report per-protocol analyses, ITT analyses are regarded as the true measure of treatment effect under the trial conditions.

10.7 Covariate Adjustment

The purpose of randomization is to achieve study groups that are comparable in every way except for the intervention of interest. Despite this, random chance may cause important prognostic factors to be distributed unequally among the study groups. Covariate adjustment can be used to reduce the effect of covariate imbalance between groups. It must be emphasized that covariate adjustment cannot be expected to completely eliminate the effect of covariate imbalance in most cases. In general, only baseline covariates measured at the study outset should be adjusted for. If the outcome is dichotomous, multivariable logistic regression can be used to obtain adjusted effect estimates. Similarly, multivariable linear regression and multivariable Cox proportional hazards model can be used to obtain adjusted effect estimates for continuous and time-to-event outcomes, respectively. Regardless of the type of regression model used, the number of covariates should not exceed 1/10 the number of patients. For example, in a study with 100 patients, no more than 10 covariates should be adjusted for. Adjustment for more variables can introduce issues with model overfitting. It is important to realize that each value of a multicategorical variable counts toward the number of covariates. For example, one might wish to adjust for a variable called “race/ethnicity” in a study where its possible values are White, Black, Hispanic, Asian, and Other. In this situation, adjusting for race/ethnicity would actually count as five covariates rather than one.

Beyond the mechanics of covariate adjustment, it is important to touch on when it should be performed in clinical trials. One school of thought is to adjust only for baseline covariates which are imbalanced between the treatment and control groups after randomization. Another approach is to decide using subject-matter knowledge which covariates have prognostic importance and to adjust only for those covariates.

In a review of trials published in 2009, about half of trials performed adjusted analyses for the main outcome as either the sole analyses (29%) or along with an unadjusted analysis (21%) [10]. In cases where an adjusted analysis was performed, the published trial and the protocol analysis plan differed in 47% of cases [10]. While there are many reasonable approaches to take with regard to whether and how covariate adjustment is performed, it is crucial that the strategy for the primary analysis be prespecified. It is often possible to change the direction and magnitude of findings by serially adjusting for different combinations of covariates. Any secondary analyses that were not prespecified should be clearly labeled as such and interpreted as less robust to bias than the prespecified analysis.

10.8 Subgroup Analyses

In subgroup analyses, the investigator looks at particular subgroups rather than the overall study cohort to examine whether different groups respond differently to the intervention. Subgroup analyses are appealing because they are in line with the goals of "precision medicine," or attempting to tailor treatments based on unique patient characteristics. One might wish to know whether an intervention in a positive trial is particularly effective in a certain subset of patients. Alternatively, in a negative trial it is appealing to attempt to define subgroups of patients where the intervention *was* effective. For these reasons, over 25% of RCTs report subgroup analyses [11]. However, many RCTs do not use best practices when conducting subgroup analyses and make claims that are not supported by their subgroup analyses [11, 12].

Subgroup analyses should be prespecified in the study protocol of the clinical trial before randomization. Investigators should include a clear hypothesis and anticipated direction of effect in each subgroup analysis. Subgroup analyses should be limited in number and adjustment for multiplicity should be considered to reduce the risk of false-positives. If many subgroup analyses are performed, it is often possible to generate findings that support one's preconceived notions and biases. Therefore, subgroup analyses that are not prespecified should be interpreted by readers as more susceptible to bias. Planning important subgroup analyses ahead of time can also help to ensure that equal numbers of participants are allocated to each treatment arm within the subgroup through stratified randomization. Failing to randomize participants equally within subgroups leaves more room for the impact of confounding between arms within subgroups. Finally, it is important to realize that the presence of a statistical effect in a subgroup does not constitute evidence of a subgroup effect. Rather, the appropriate approach of establishing a subgroup test is a formal test of interaction [13]. Despite this, only a minority of trials reporting a statistically significant subgroup analysis perform a formal test of interaction [12].

10.9 Handling Missing Data

In most clinical trials, some participants will have missing data or data that is of such poor quality that it must be handled as missing. It must be emphasized at the outset that the best way to handle missing data is to minimize its occurrence during trial design and conduct. Techniques used to deal with missing data include different imputation methods such as last observation carried forward, complete case analysis, and mean or median value imputation and multiple imputation (MI) techniques. Choice of the correct method for dealing with missing data depends on the pattern of data missingness.

10.10 Patterns of Missing Data

Data can be missing in three main patterns: missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR) [14]. Data are considered MCAR if the probability of data being missing is unrelated to both observed and unobserved patient observations. The MCAR assumption is the least plausible and is rarely met. Yet if the MCAR assumption is met, then a complete case analysis will yield unbiased results. Complete case analysis refers to exclusion of any observations with missing data. Most common statistical software packages perform complete case analysis by default if missing data are present. Complete case analysis will also decrease power by decreasing sample size. If complete case analysis is used and the MCAR assumption is not met (a likely event), then estimates will be biased and the direction of the bias will be unpredictable [14].

The MAR or "ignorable" mechanism assumes that observed values can be used to predict what missing values would be [14, 15]. This assumption is more realistic than MCAR, and MI techniques must make this assumption. Finally, a pattern of MNAR occurs when missing values are dependent on unobserved or unknown factors. When MNAR has occurred, no statistical method can account for missing information.

10.10.1 Single Imputation Methods

Before sophisticated imputation schemes were available, it was common to use single imputation method such as carrying the last observation forward or replacing the missing value with a likely value such as the mean, median, or mode. These approaches artificially decrease variance since many observations will have a single value. Decreasing variance tends to artificially increase precision, creating spuriously small p-values and spuriously tight confidence intervals.

10.10.2 Multiple Imputation Methods

Better point estimates and measures of uncertainty can be obtained using hot deck imputation and multiple imputation (MI). In hot deck imputation, each missing value is replaced with the value from the most similar case for which the variable is not missing. However, this technique performs poorly when many observations have a missing data point [16].

MI is a technique in which missing values are generated by creating plausible numbers based on distributions of and relationships among observed variables in the dataset [17]. In comparison with single imputation methods discussed above, missing data are filled in many times using MI techniques, generating numerous plausible values for each missing value. This process is completed in two stages. First, replacement values ("imputations") are generated, resulting in many datasets with replaced missing information. The number of missing datasets is set by the analyst. There are no absolute rules for how many imputed datasets to use, but more are generally better. Graham et al. showed that the power to detect a small difference in outcome falls off dramatically with small numbers of imputed datasets [18]. They argue for using at least 40 imputed datasets when 50% of observations have some missing data to guarantee less than a 1% power falloff compared to an analysis with no missing data. After stage 1 is completed, the intended analysis (*t*-test, regression, etc.) is conducted within each imputed dataset. Finally, the treatment effect estimates of interest from each imputed dataset are combined. The reported standard errors and confidence intervals allow for uncertainty due to missing data [17].

10.11 Other Measures Used in Clinical Trials

There are other measures used to evaluate the results in clinical trials, including the number needed to treat (NNT) and the fragility index (FI). The NNT is calculated as the reciprocal of the absolute risk reduction and is an aggregate measure of the benefit of a treatment that represents the number of patients who would need to be treated to prevent one event (outcome) [19]. It is important to understand that the NNT varies inversely with baseline risk [20]. Therefore, the NNT rarely appears favorable in low-risk populations even if the intervention is highly efficacious. The NNT cannot be calculated for continuous outcomes, and it is best suited for dichotomous outcomes with short follow-up time. In the case of a dichotomous time-to-event outcome (i.e., survival), it is unlikely that the NNT will be constant over time given the increasing importance of competing risks as time passes [20].

The FI is a measure of the robustness (or fragility) of the results of an RCT. The FI is the number of patients whose status would change from nonevent (not experiencing the primary outcome) to an event to make the study lose statistical significance [21]. The FI is a measure of how many events the statistical significance reported in a trial depends on; a smaller FI corresponds to a more fragile trial result.

Online calculators for FI calculation are available [22]. The FI can be used for dichotomous outcomes but not continuous outcomes. It can be applied to time-to-event binary outcomes, but it may be inappropriate to apply it to situations where the number of events in each group is similar but differs in timing (i.e., a situation where patients in both arms die eventually, but those in the intervention arm live longer). Most surgical and trauma trials have a low FI [21].

10.12 Reporting the Results from Clinical Trials

Reporting the results of a clinical trial is one of the most critical responsibilities of a surgical investigator. It is important to provide all necessary information about external and internal validity that will allow a reader to evaluate the study's conclusion. Providing comprehensive and complete data allows the reader to determine whether the study results are valid and can be generalized to patients they care for in their own clinical practice. Moreover, standardized reporting allows investigators to compare findings with other published studies and facilitate downstream synthesis.

In order to attain the highest standards for reporting clinical trial results, the Consolidated Standards of Reporting Trials (CONSORT) criteria were established in 1996, and they have subsequently been revised in 2001 and 2010 [23, 24]. The CONSORT criteria were developed to specifically improve the quality of research reporting of randomized controlled trials and are now accepted as general guidelines for investigators who publish the results from clinical trials. Currently, most leading medical journals require that the CONSORT 25-item checklist be used along with a flow diagram to show the reader how the clinical trial was designed, analyzed, and interpreted (available at http://www.consort-statement.org).

Broader reporting issues in clinical trials have also been addressed by the International Committee of Medical Journal Editors (ICMJE). The ICMJE has released guidelines that cover ethical principles related to reporting and public registration of clinical trials. Specifically, ICMJE recommends that all medical journal editors require the registration of clinical trials in a public trial registry at or before the time the first patient is enrolled as a condition of consideration for publication. This includes registration in the ClinicalTrials.gov registry or the WHO International Clinical Trials Registry Platform (ICTRP). The ICMJE also encourages investigators to update these registry sites with the full journal citation when clinical trial results are published.

One of the main purposes of clinical trial registration is to prevent selective publication and selective reporting of research outcomes. Publication bias refers to the tendency for only studies with "positive" findings to be selected for publication, whereas clinical trials with "negative" results may never get published. Other benefits of requiring clinical trials to be publicly registered include preventing unnecessary duplication of research efforts and helping patients and the public know what trials are planned or ongoing into which they might want to enroll.

10.13 Conclusion

Prospective randomized controlled trials provide the best evidence for deciding the value of surgical interventions and can have a direct impact on patient care. However, because most surgical interventions are complex and multifactorial, surgical clinical trials pose special challenges for the investigator in terms of design, analysis, and reporting. It is critical for surgical investigators to have a clear understanding of clinical trials methodology and tools needed to conduct and interpret clinical trials. This chapter provides an overview of how to handle data at different stages of a clinical trial, including common methods used in the process of randomization, statistical techniques applied in the analysis, and how to report the trial results in a manner that allows readers to accurately interpret the study findings.

References

1. Amrhein V, Greenland S, McShane B. Scientists rise up against statistical significance. Nature. 2019;567(7748):305–7.
2. Ionnidis JPA. Retiring statistical significance would give bias a free pass. Nature. 2019;567(7749):461.
3. Ioannidis JPA. The proposal to lower P value thresholds to .005. JAMA. 2018;319(14):1429–30.
4. Chavalarias D, Wallach JD, Li AHT, Ioannidis JPA. Evolution of reporting P values in the biomedical literature, 1990-2015. JAMA. 2016;315(11)
5. Norton EC, Dowd BE, Maciejewski ML. Odds ratios: current best practices and use. JAMA. 2018;320(1):84–5.
6. Norton EC, Dowd BE. Log odds and the interpretation of logit models. Health Serv Res. 2018;53(2):859–78.
7. Norton EC, Miller MM, Kleinman LC. Computing adjusted risk ratios and risk differences in Stata. Stata J. 2013;13(3):492–509.
8. Uno H, Claggett B, Tian L, et al. Moving beyond the hazard ratio in quantifying the between-group difference in survival analysis. J Clin Oncol. 2014;32(22):2380–5.
9. Weir IR, Marshall GD, Schneider JI, et al. Interpretation of time-to-event outcomes in randomized trials: an online randomized experiment. Ann Oncol. 2019;30(1):96–102.
10. Saquib N, Saquib J, Ioannidis JP. Practices and impact of primary outcome adjustment in randomized controlled trials: meta-epidemiologic study. BMJ. 2013;347:f4313.
11. Kasenda B, Schandelmaier S, Sun X, et al. Subgroup analyses in randomised controlled trials: cohort study on trial protocols and journal publications. BMJ. 2014;g4539:349.
12. Wallach JD, Sullivan PG, Trepanowski JF, Sainani KL, Steyerberg EW, Ioannidis JP. Evaluation of evidence of statistical support and corroboration of subgroup claims in randomized clinical trials. JAMA Intern Med. 2017;177(4):554–60.
13. Oxman AD, Guyatt GH. A consumer's guide to subgroup analyses. Ann Intern Med. 1992;116(1):78–884.
14. Newgard GD, Lewis RJ. Missing data: how to best account for what is not known. JAMA. 2015;314(9):940–1.
15. Little RJA, Rubin DB. Statistical analysis with missing data. 2nd ed. Wiley: Princeton, NJ; 2002.
16. Royston P. Multiple imputation of missing values. Stata J. 2004;4(3):227–41.
17. Peng L, Stuart EA, Allison DB. Multiple imputation: a flexible tool for handling missing data. JAMA. 2015;314(18):1966–7.

18. Graham JW, Olchowski AE, Gilreath TD. How many imputations are really needed? Some practical clarifications of multiple imputation theory. Prev Sci. 2007;8:206–13.
19. Laupacis A, Sackett DL, Roberts RS. An assessment of clinically useful measures of the consequences of treatment. N Engl J Med. 1988;318:1728–33.
20. McAlister FA. The "number needed to treat" turns 20—and continues to be used and misused. CMAJ. 2008;179(6):549–53.
21. Tignanelli CJ, Napolitano LM. The fragility index in randomized clinical trials as a means of optimizing patient care. JAMA Surg. 2018.
22. ClinCalc LLC. Fragility index calculator. http://clincalc.com/Stats/FragilityIndex.aspx. Published 2018. Accessed 2 Apr.
23. Schulz KF, Altman DG, Moher D, Guyatt GH. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ. 2010;340:c332.
24. Begg C, Cho M, Eastwood S, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. JAMA. 1996;276(8):637–9.

Chapter 11
Data Safety Monitoring Boards

Rachael A. Callcut

11.1 Data Safety Monitoring Boards (DSMBs)

Data Safety Monitoring Boards (DSMBs) can also be known as data monitoring committees or data review boards. The overall purpose of the DSMB is to insure the integrity and safety of clinical trial research [1–3]. Although no universally accepted definition exists for these committees, the guidance issued by the Federal Drug Administration (FDA) in 2006 is the most widely utilized to frame the scope of a DSMB [3]. The FDA notes that the DSMB is a group of individuals that possess pertinent scientific expertise that review, on a regular basis, the interim research data from ongoing clinical trials [4, 5]. The group serves to advise the sponsor and/or researchers of the study on the safety of the trial, the continuing validity of the trial, and the scientific merit of the trial. Although DSMBs are now commonly utilized, there remains no singular entity responsible for oversight of these groups, and the operationalization of these committees can be variable.

11.2 History

As research has evolved in the last 50 years, the need for monitoring the safety of clinical trials has increased [1, 6]. The origins of trial oversight began in the early 1960s when the National Institutes of Health (NIH) began funding multicenter clinical trials. A taskforce was led by Bernard Greenberg on behalf of the NIH (former National Heart Institute) and delivered their report in 1967 [6, 7]. This report is largely credited as the origin of data monitoring committees [4]. Amongst the most important recommendation was the call for establishment of independent groups of

R. A. Callcut (✉)
Department of Surgery, University of California San Francisco, San Francisco, CA, USA
e-mail: Rachael.Callcut@ucsf.edu

T. M. Pawlik, J. A. Sosa (eds.), *Clinical Trials*, Success in Academic Surgery,
https://doi.org/10.1007/978-3-030-35488-6_11

experts to advise the institute on the conduct of a trial [7]. Specifically, all group members were to have no direct involvement in the trial under surveillance. Interestingly, although this report was submitted in 1967, it was not published in the scientific literature until 1988 [6].

Following the publication in 1988, industry-sponsored trials, specifically those in the pharmaceutical domain, began to increase the use of DSMBs. The NIH also increased their guidance on DSMBs in the 1990s with a specific mandate in 1994 that a data safety monitoring plan was required for every clinical trial funding with federal funds. In 1998, it became mandated that all federal funded, multicenter trials have a DSMB [3, 8]. Beginning in the early 2000s, the adoption of data monitoring committees became common place for industry trials [4]. In 2005, DAMOCELS study group proposed a charter for data monitoring boards [5, 9], and in 2006, the FDA issued a formal guidance for Clinical Trials which remains in place today [4].

Currently, the NIH requires all Phase-III multicenter trials and blinded trials to have a DSMB if the study involves any risk [3]. Importantly, the details of the individual institute policies vary [7, 8]. It is now generally accepted that this includes randomized interventional control trials (treatment versus placebo, or comparing two treatments), high risk studies (due to safety concerns), and studies of early novel therapies (where safety data are lacking), even if conducted without federal funding [7]. Importantly, even if a DSMB is not required, all NIH funded clinical trials require a data safety monitoring plan [8]. It is advisable to consult with your local IRB for further guidance on when a DSMB versus only a data safety monitoring plan is needed.

11.3 Objectives of a DSMB

DSMBs are a fundamental component of insuring that clinical trials are conducted safely, ethically, and remain scientifically sound. The group functions in an independent, advisory capacity to the study sponsors or researchers [2]. They have the difficult task of insuring patient safety while also maintaining as much scientific validity of a study is possible. First and foremost, safety is the highest priority. These groups also play a pivotal role in optimizing the length of a trial. As data in a trial are accumulated, the DSMB conducts interim reviews. The groups recommend early termination for both trials reaching futility and also for trials where one intervention is far superior, and thus, further enrollment is viewed as unethical. The committees review protocol violations, drop-out rates, and conduct interim analysis of trial data.

11.4 DSMB Charter

The objectives, organization, and expectations for the DSMB are often defined through a “DSMB Charter.” The charter should be drafted prior to the start of a trial to stipulate the DSMB membership, roles and responsibilities, meeting timeframes,

safety monitoring plan, data analysis plan, interim stopping criteria, conflict of interest procedures, communication plan, and a confidentiality statement [1, 2, 6, 8, 10]. The charter also stipulates voting procedures. Although some committees use a formal voting process, it is preferred for the committees to reach consensus [2]. Importantly, the criterion for initiating an unplanned review and decision of what would prompt complete unblinding (if applicable) of data should be determined in advance [8].

Ideally, these documents should also inform the DSMB members of their legal risk in participating in the committee and if they will be indemnified by the sponsor. DSMB members have been called upon to testify in legal matters pertaining to trial risks from study participants in the past [2, 6]. In most cases of industry-sponsored research trials, the sponsor will represent and indemnify the members from any personal liability as long as they conducted themselves within the law and ethics of their role [6]. However, federally sponsored trials do not contain this indemnification, and potential participants in DSMBs should consider the implications of this [2]. DSMB members usually sign a contract to be on the panel, and members can negotiate in these contracts to be indemnified [2].

11.5 DSMB Membership and Training

Prior to initiating a study, the DSMB is selected. There is no standard size or composition for a committee. Typically, the groups include a minimum of 3–7 members [1], but can be much larger for complicated trials. Committee members are selected usually by either the sponsor or the principal investigator [4]. The expertise of the DSMB should include at a minimum an experienced biostatistician, ideally with a clinical trial background, a clinical expert in the field under investigation, and at least one other scientist [1, 6, 8]. Often these committees have a medical ethicist [8] and can even have a patient advocate [1]. Best practices include having a committee with gender and ethnic diversity [4]. Members must have no conflict of interest or involvement with any person, organization, institution, or sponsor of the trial [2, 8]. They may not participate in any other aspect of the study and must agree to maintain confidentiality at all times [2].

There is no formal training for members of DSMBs, and in a recent survey, only 8% of data monitoring committee members reported being formally trained [2]. Historically, when the numbers of DSMBs needed were low, the average experience of the members was robust. However, with the increased regulatory requirements and number of trials growing, the demand for the services of those willing to serve on a DSMB has outpaced those available [2, 3]. This has become a concern amongst the veteran clinical trial experts [6]. The time commitments, especially for larger or more complex trials, can be significant [7]. Experts have called for an increase in the acknowledgment of the members (if the individual desires) of the committees in formal publications [7].

11.6 DSMB Procedures

DSMBs, especially for larger clinical trials, have a series of pre-specified meetings, with at least the final meeting occurring in person. The first meeting is an *organizational meeting* [4]. The charter is reviewed at this meeting with focus on the roles and responsibilities of members, protocol safety monitoring design, and the statistical analysis plan. Often at the first meeting, both the sponsor and study investigators are present. This provides an opportunity for the DSMB members to identify and address key flaws prior to study enrollment.

Once enrollment has commenced, most DSMBs will conduct an *early safety review*. This review can help identify early issues with enrollment, early protocol violations indicative of study procedures needing clarification, and address quality control issues [11]. Subsequent meetings tend to coincide with the planned *interim analyses*. The number of interim analyses and study stopping criteria are predetermined and stipulated in the charter. Interim analyses are typically conducted with the DSMB knowing which patients are in treatment group A or B [1]. This allows them to render an independent decision. Although a committee will sometimes know assignment of a patient to treatment A or B, their report of the meeting should only identify the groups as A or B [6]. The DSMB may or may not know what treatment arm is A or B and, if necessary, can request complete unblinding of data to assist in assessing safety of a trial [5]. However, the FDA advocates that the DSMB have the unblended data [4]. The interim meetings are also focused on the conduct of the study including reviewing serious adverse event reports and individual center performance [3].

A progress report from the principal investigator is commonly provided to the board, and before rendering any recommendations, the committee is tasked with determining if the information provided is sufficient to adequately determine safety and welfare for the study participants [6]. The board also reviews the efficacy, to date, of the study. Although often not stated, the group has to also consider if there has been an advance in the field outside of the on-going trial that would impact the results of the trial [4]. This information could be the reporting of a similar study showing lack of efficacy or clear benefit, thus the trial being conducted no longer meets the standards of equipoise.

At each step of the oversight, the DSMB can make a variety of recommendations. For safety issues, these can focus on modifications to study protocols, early termination of a part or all of a study, or corrective action for one or more study sites [6, 12]. They also have the difficult task of making recommendations to sponsors and regulatory agencies about early termination of a study [3]. This can occur in a favorable direction because the study endpoints have been reached with a more pronounced beneficial effect than the study was originally powered for, thus further patient enrollment is unlikely to change the study results. The DSMB has to consider "the clinically important difference or the minimum magnitude of treatment benefit large enough to offset the treatment harms [10]." This requires the use of Baysesian probability statistical analysis and not just simple statistical analysis [10].

Alternatively, and often more controversial, they can determine if a study should cease on the grounds of futility [11, 12]. The decision to stop a trial for futility is a difficult decision. The charter should list in advance the specific criteria for ending a trial early [6, 10]. The threshold is generally met when a significant difference between the treatment groups is unlikely to be identified based on the interim analysis [10]. In other words, the trial data will be unlikely to reject a null hypothesis [13].

All DSMBs conduct a *final close-out* meeting. This occurs either at the planned termination of the study or following a decision to stop a trial early. The committee will generate a final report with summary recommendations to the study sponsor. Although the DSMB functions independently of the Institutional Review Boards where the trial is being performed, adverse event reports are to be shared with the IRB and DSMB collectively [3]. The DSMB functions as an overall safety monitor for all sites, whereas institutional IRBs have oversight of a single site. In addition, DSMBs may have interactions with groups such as the FDA for trials under FDA guidelines and independent trial medical monitors.

11.7 Data Safety Monitoring Plans

A key function of the oversight committee for a trial is to insure that the data safety monitoring plan is adequate prior to trial commencement and adhered to throughout the trial. Although there is no standard format for a safety plan, there are fundamental principles to consider when conducting a trial. The plan should clearly state how the trial progress will be monitored. Risk to participants, how participants will be protected from harm, how safety will be assessed for participants, the steps to report an unusual or adverse event, how data accuracy will be checked, and a management strategy for conflicts of interest should also be included. Protocol compliance should also be defined. It must also include what information will be reviewed at interim analysis, the timing of the interim analyses, stoppage criteria, and communication procedures for multicenter studies. Monitoring plans should be utilized in all clinical trial research even if a DSMB is not required [11]. Federal regulations require such monitoring (45 CFR 46.111(a)(6); 21 CF 56.111(a)(6)) for minimal risk or greater than minimal risk studies [8].

References

1. Calis KA, Archdeacon P, Bain RP, Forrest A, Perlmutter J, DeMets DL. Understanding the functions and operations of data monitoring committees: Survey and focus group findings. Clin Trials. 2017;14(1):59–66.
2. Fleming TR, DeMets DL, Roe MT, Wittes J, Calis KA, Vora AN, Meisel A, Bain RP, Konstam MA, Pencina MJ, Gordon DJ, Mahaffey KW, Hennekens CH, Neaton JD, Pearson GD, Andersson TLG, Pfeffer MA, Ellenberg SS. Data Monitoring Committee: promoting best practices to address emerging challenges. Clin Trials. 2017;14(2):115–23.

3. Levinson DR. Data safety and monitoring boards in NIH clinical trials: meeting guidance, but facing some issues. Washington, DC: Department of Health and Human Services, Office of the Inspector General; 2013. (OEI-12-11-00070).
4. FDA. Guidance for clinical trial sponsors: establishment and operations of clinical trial monitoring committees. 2006. www.fda.gov. Accessed 28 Mar 2019.
5. Marston L, McKenzie DR, Freemantle N. Pursuit of truly independent data monitoring committees in research. BMJ. 2016;353:i2309.
6. DeMets DL, Ellenberg SS. Data monitoring committees – expect the unexpected. NEJM. 2016;375(14):1365–71.
7. Gewandter JS, Kitt RA, Hunsinger MR, Poku J, Lozano J, Chaudari J, Evans S, Gross RA, McDermott MP, Rowbothan MC, Turk DC, Dworkin RH. Reporting of data monitoring boards in publications of randomized clinical trials is often deficient: an ACTTIION systematic review. J Clin Epidemiol. 2017;83:101–7.
8. Bierer B, Winkler S, Bernstein H, Myerson J, Kim M. Data safety monitoring guidance. Harvard Catalyst, Regulatory Foundations, Ethics, and Law Program. Boston, MA: The Harvard Clinical Translational Science Center; 2018.
9. DAMOCLES Study Group. A proposed charter for clinical trial data monitoring committees: helping them do their job well. Lancet. 2005;365(9460):711–22.
10. Tyson JE, Pedroza C, Wallace D, D'Angio C, Bell EF, Das A. Stopping guidelines for an effectiveness trial: what should the protocol specify? Trials. 2016;17(240):1–4.
11. Ellenberg SS, Culbertson R, Gillen DL, Goodman S, Schrandt S, Zirkle M. Clin Trials. 2015;12(5):530–6.
12. Swedberg K, Borer JS, Pitt B, Pocock S, Rouleau J. Challenges to data monitoring committees when regulatory authorities intervene. NEJM. 2016;374(16):1580–4.
13. Harrington D, Drazen JM. Learning form a trial stopped by a Data and Safety Monitoring Board. NEJM. 2018;378(21):2031–2.

Chapter 12
Planning for Data Monitoring and Audits

Benjamin K. Poulose

12.1 Good Clinical Practice Guidelines

Standards regarding data monitoring and auditing have been well established in the setting of clinical trials. The three pillars of data collection are summarized in Fig. 12.1. These standards, developed through the International Council for Harmonization (ICH) of Technical Requirements for Pharmaceuticals for Human Use, have most recently been summarized by the United States Food and Drug Administration (FDA) in 2018 as part of recommended Good Clinical Practice [1] (Table 12.1).

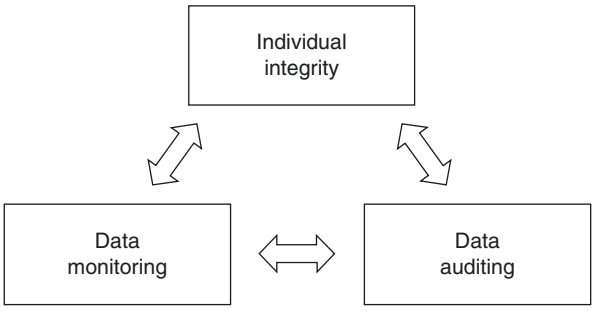


Fig. 12.1 The three pillars of data collection: individual integrity in recording and gathering of information is vital to seeking truth

B. K. Poulose (✉)
Division of General and Gastrointestinal Surgery, Center for Abdominal Core Health, The Ohio State University Wexner Medical Center, Columbus, OH, USA
e-mail: benjamin.poulose@osumc.edu

T. M. Pawlik, J. A. Sosa (eds.), *Clinical Trials*, Success in Academic Surgery,
https://doi.org/10.1007/978-3-030-35488-6_12

Table 12.1 Core principles of International Council for Harmonization's Good Clinical Practices [2]

1. Clinical trials should be conducted in accordance with ethical principles that have their origin in the Declaration of Helsinki and that are consistent with GCP and the applicable regulatory requirement(s)
2. Before a trial is initiated, foreseeable risks and inconveniences should be weighed against anticipated benefit for the individual trial subject and society. A trial should be initiated and continued only if the anticipated benefits justify the risks
3. The rights, safety, and well-being of the trial subjects are the most important considerations and should prevail over the interest of science and society
4. The available non-clinical and clinical information on an investigational product should be adequate to support the proposed clinical trial
5. Clinical trials should be scientifically sound and described in clear, detailed protocol
6. A trial should be conducted in compliance with the protocol that has received prior institutional review board (IRB)/independent ethics committee (IEC) approval/favorable opinion
7. The medical care given to and medical decisions made on behalf of subjects should always be the responsibility of a qualified physician or, when appropriate, of a qualified dentist
8. Each individual involved in conducting a trial should be qualified by education, training, and experience to perform his or her respective task(s)
9. Freely given informed consent should be obtained from every subject prior to clinical trial participation
10. All clinical trial information should be recorded, handled, and stored in a way that allows its accurate reporting, interpretation, and verification
11. The confidentiality of records that could identify subjects should be protected, respecting the privacy and confidentiality rules in accordance with the applicable regulatory requirement(s)
12. Investigational products should be manufactured, handled, and stored in accordance with applicable Good Manufacturing Practice (GMP). They should be used in accordance with the approved protocol
13. Systems with procedures that assure the quality of every aspect of the trial should be implemented

The core principles 10, 11, and 13 serve as the cornerstone to ensure integrity of data collection. These principles primarily apply to the performance of clinical trials, but many concepts have also been adopted in data collection for registries utilized for research and quality improvement.

12.2 Data Monitoring

Data monitoring refers to the ongoing act of overseeing clinical trial data accrual or data gathering for other reasons (e.g., quality improvement). This process occurs during the routine procedures specified at the start of data collection to ensure compliance with those procedures. In high stakes clinical trials, a formal data safety monitoring board (DSMB) or data safety monitoring committee (DSMC) is also

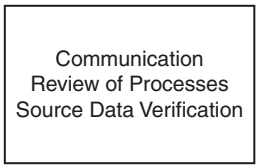


Fig. 12.2 Goals of data monitoring

created to periodically check compliance with protocols, assess for safety events, and halt data collection efforts if needed as a trial proceeds to completion.

One of the best summaries of common data monitoring practices for research is provided by the FDA [3]. The three goals of monitoring include (1) communication with study site staff, (2) review of site processes, procedures, and records, and (3) source data verification (Fig. 12.2). In general, monitoring methods are divided into onsite or centralized processes. On-site monitoring is performed in person at the site where the research or data gathering is being performed. If a sponsor is involved, onsite monitoring has the benefit of minimizing communication gaps between sponsor and site investigator and helps build a relationship between the two. On the other hand, onsite monitoring is time- and personnel-intensive.

Centralized monitoring offers the advantages of increased efficiency and less time involvement, especially if electronic health records are utilized. The overall trend in clinical trials is toward centralized monitoring due to its greater efficiency. Personnel need to be well trained in data security and privacy practices to minimize risk involved in the transfer of large amounts of protected health information. The frequency and depth of monitoring procedures varies greatly on the rigor required, resources available, and intended use of the information collected.

Alternative monitoring techniques have emerged in a further effort to increase efficiency and reduce the burden of monitoring. A targeted approach can be risk-based or statistically based. In a risk-based approach, sites or individuals deemed high risk are monitored more closely. High risk is defined as sites or individuals known or suspected to have protocol violations, errors in data collection, or even data fabrication. In these high risk situations, it is vital to determine if the error is in source documentation (e.g., not documenting complications in the health record) or in completion of data reporting forms. The latter is usually readily identified; the former can be very difficult to ascertain. Risk-based approaches can also focus data monitoring on key variables in a study or process (exposure, outcome, important known confounders). Appropriate use of a risk-based approach is also dependent on the investigators and/or sponsor. Inappropriately targeting sites, problematic variables, or biased scrutiny of potentially negative outcomes should be strictly avoided. These issues can be subtle to detect and very difficult to address when found. Statistically based approaches can be used to monitor particular data points, individuals, or sites, and alert the data monitoring team to potential issues. These techniques are often used by formal DSMB or DSMC groups overseeing clinical trials. Additionally, statistical methods can be used to detect unusual data distributions, outliers, data completeness, and unexpected

variance (too much or too little). These methods require collaboration with experienced biostatisticians to minimize bias and appropriately account for the monitoring process itself when needed.

12.3 Data Auditing

Data auditing, in contrast to data monitoring, usually involves a third party review of regulatory procedures and compliance with laws and standards. Unlike routine operations, such as data monitoring, data auditing usually involves unannounced reviews that are often fraught with anxiety, immediate scheduling requests, and general consternation. This can have a negative impact on teams responsible for gathering data, especially if serious deficiencies are found or if regulatory lapses are discovered. The request for an audit alone can have negative consequences for investigators. This is unfortunate as any entity (even competing institutions or companies) can trigger an unannounced audit from regulatory bodies. Institutional leadership should recognize this possibility, especially in an increasingly competitive fiscal and academic environment. The best antidote to these situations is adherence to protocols and procedures, care in documentation, and accommodation of the auditing team. Weiss et al. summarize audit preparation concisely into three main categories which are summarized in Fig. 12.3 [4].

Review of individual patients is a key component to a data audit and involves predictable items (Table 12.2).

The logistical aspects of performing the audit are also important. The identity of the auditors should be voluntarily provided and confirmed along with necessary documentation surrounding the audit. Institutional leadership should be immediately notified. It is extremely helpful if an institutional representative with experience in audits assists with collation and presentation of information, even if that person does not have direct involvement with the study or process being audited. A central location

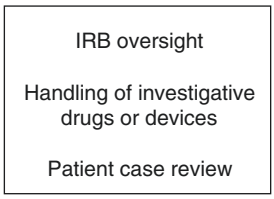


Fig. 12.3 Important components of data auditing; IRB (Institutional Review Board)

Table 12.2 Patient case review items involved in data audits; adapted after Weiss et al. [4]

1. Consent form signing, dating it, and completion
2. Protocol eligibility
3. Protocol-directed treatment
4. Verification of treatment response
5. Adverse event recording
6. Accuracy of data recording and submission

should be provided to the auditing team and key personnel made readily available. The importance of the latter cannot be emphasized enough. Clinical, administrative, and personal schedules will need to be altered during this process to show the auditors good faith in completing the audit. Keeping meticulous records greatly facilitates review. Well-organized, physical binders are best used for auditing even if records are maintained electronically. If this is the case, key documents should be printed, organized, and separate binders presented to auditors for each patient. In addition, a computer should be available to cross-check items as needed.

The single most anxiety-provoking element of unannounced audits, especially from the FDA, is not knowing the reason for the audit throughout the process. Teams should rely on the basics: adherence to protocols and procedures, care in documentation, and having the highest integrity and organization throughout data collection. Much comfort is gained in these difficult situations knowing that proper protocols and procedures have been followed, care has been taken in documentation, and teams act with highest levels of integrity.

12.4 Data Monitoring and Auditing in Practice

There are several common scenarios where surgeons can experience various aspects of data monitoring and auditing. The following examples illustrate these concepts in practice.

12.4.1 Investigator Initiated Studies

Investigator initiated studies (typically unfunded) comprise the vast majority of studies in the surgical literature. There are no standards for data monitoring, and data auditing is extremely rare. These studies rely solely on investigator integrity for truthfulness in data.

12.4.2 Clinical Trials

Clinical trials range from industry-sponsored studies to multicenter, federally funded randomized trials. Data monitoring processes are typically robust and include error-checking mechanisms with data entry, confirmation of processes and protocols, and source document verification. A sample of data is usually monitored (10–20%) using a risk-based or statistically based approach. For higher stakes trials (e.g. Investigational Device Exemption trials), DSMB or DSMC groups are used and up to 80% of data are reviewed. Audits are more commonplace, especially with higher stakes industry-sponsored trials.

12.4.3 Clinical Registries

Clinical registries have seen an increase in utilization to meet a gap in knowledge not provided by evaluation of the medical record alone. Several groups have excellent recommendations on standards for registry inception, maintenance, and governance [5–7]. Having a data monitoring strategy is critical for the integrity of a registry. Registries can be established for research purposes and also for routine healthcare operations such as quality improvement. Data monitoring in the context of registries typically involves a completion and accuracy assessment. For data completion, a determination is made to ensure that the intended types of patients are actually being entered into the registry without bias in case entry. This can be performed by comparing billing records to records entered into the registry itself over a specified period of type for particular physicians. Data accuracy is performed by manual record review, comparing data entered into the registry versus the health record. This can be performed in a targeted fashion using either a risk-based or statistical approach to increase efficiency. Typically, 3–10% of records are reviewed on a rolling basis with the intent to review each site at least once over a given time period. An interesting conundrum has evolved as data are often entered into a detailed clinical registry without necessarily requiring the same information be entered into a patient's health record. New ways of interpreting "source documents" should be sought in light of this trend.

12.4.4 Registry-Based Clinical Trials

Registry-based clinical trials offer an innovative and efficient way of performing clinical trial while greatly reducing the cost and inefficiencies of traditional clinical trials. The cornerstone of these types of trials is a robust registry infrastructure over which the clinical trial can be "overlaid." In these situations, each participating site is usually required to have undergone a data assurance review (involving completeness and accuracy) within the past year to maintain registry standards. Depending on the type of trial, a standard level of record review for clinical monitoring is then established.

12.5 Investigator and Clinician Integrity

The key element to successful and truthful data collection and presentation rests with investigators and clinicians. Great care should be taken in clinical documentation, especially in the era of electronic health records where errors can readily be propagated across several documents. The overwhelming tendency is to document the minimum necessary for billing and legal purposes; this needs to be balanced

against the need for more comprehensive documentation used as a source document for research and quality improvement. No amount of data monitoring or data auditing can replace investigators, clinicians, and teams committed to gathering truthful information while adhering to robust standards and protocols.

References

1. E6(R2) (2018) Good clinical practice: integrated addendum to ICH E6(R1) guidance for industry E6(R2) good clinical practice: integrated addendum to ICH E6(R1) guidance for industry procedural contains nonbinding recommendations. OMB control No. 0910-0843 Expiration Date 09/30/2020.
2. Vijayananthan A, Nawawi O. The importance of good clinical practice guidelines and its role in clinical trials. Biomed Imaging Interv J. 2008;4:1–4.
3. FDA. Guidance for industry oversight of clinical investigations—a risk-based approach to monitoring guidance for industry oversight of clinical investigations—a risk-based approach to monitoring. Clin Investig. 2011;2019. OMB control No. 0910-0843 Expiration Date 09/30/2020.
4. Weiss RB, Tuttle SS. Preparing for clinical trial data audits. J Oncol Pract. 2006;2:157.
5. PCORI. Category 6: standards for data registries. https://www.pcori.org/research-results/research-methodology/methodology-standards-academic-curriculum/category-6-standards. Accessed 29 Mar 2019.
6. The Pew Charitable Trusts. Next steps to encourage adoption of data standards for clinical registries. https://www.pewtrusts.org/en/research-and-analysis/fact-sheets/2016/11/next-steps-to-encourage-adoption-of-data-standards-for-clinical-registries. Accessed 29 Mar 2019.
7. Registries for evaluating patient outcomes: a user's guide: 3rd edition. Effective Health Care Program. https://effectivehealthcare.ahrq.gov/topics/registries-guide-3rd-edition/research. Accessed 29 Mar 2019.

Chapter 13
The Budget

Shuab Omer and Faisal G. Bakaeen

13.1 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]. For the fiscal year 2020, the President has proposed the budget for National Institute of Health (NIH) of just $ 34.4 billion for biomedical research which has essentially remained approximately the same since 2017 [3]. 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 [4–6]. In 2018, the NIH received 28,072 applications for an RO1 grant out of which just 5003 were approved corresponding to a success rate of just 17% [7].

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 [5]. This dismal situation for NIH funding contrasts strikingly with all the new emerging avenues for research that are avail-

S. Omer
Advanced Cardiopulomary Therapies and Transplantation, University of Texas Health Science Center, The McGovern Medical School, Houston, Teaxs, USA

F. G. Bakaeen (✉)
Department of Thoracic and Cardiovascular Surgery, The Cleveland Clinic Heart and Vascular Institute, Cleveland, OH, USA
e-mail: fbakaeen@bcm.edu

T. M. Pawlik, J. A. Sosa (eds.), *Clinical Trials*, Success in Academic Surgery,
https://doi.org/10.1007/978-3-030-35488-6_13

able now due to rapid advances in proteomics, genetic sequencing, stem cells, and other technological advances [4–6].

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.

13.2 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. 13.1). For multi-institutional study applications, a separate budget must be submitted for each participating site.

13.3 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 [8]. 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 [9].

13.3.1 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 [9]. Before embarking on any project, carefully read the funding opportunity announcement for budget criteria and formulate your budget accordingly.

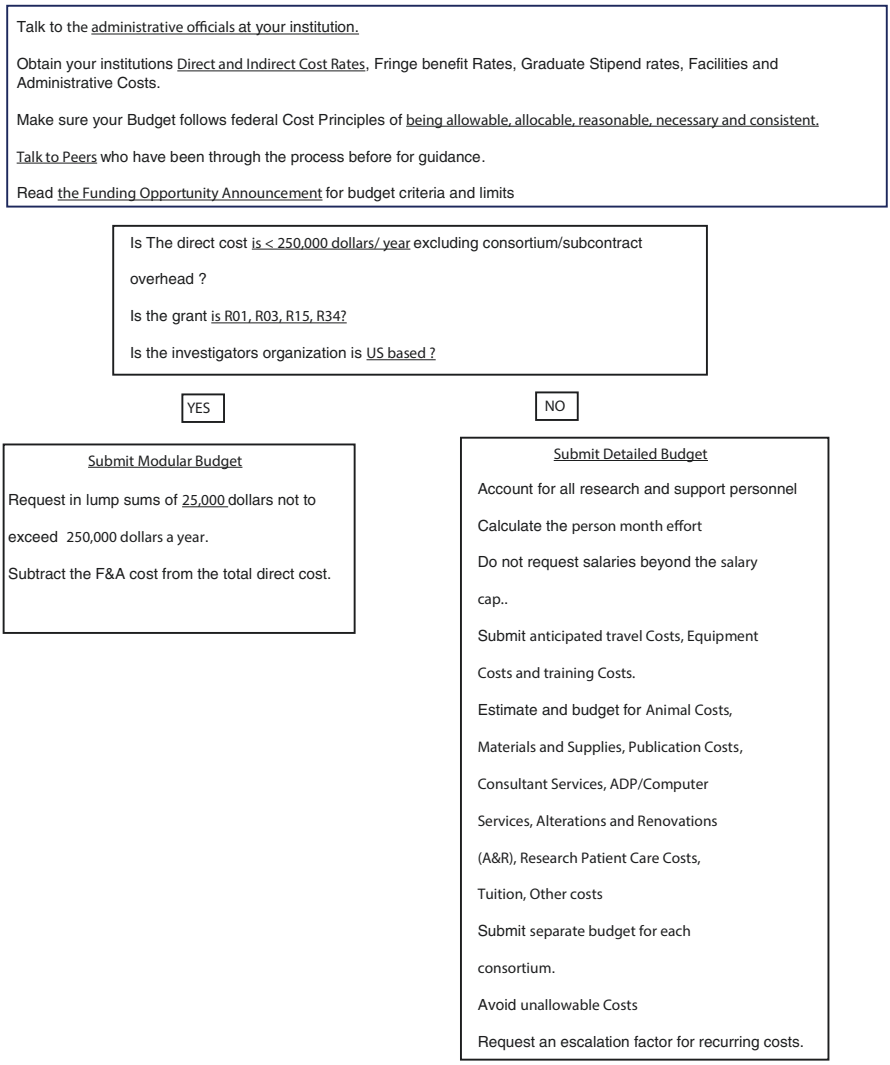


Fig. 13.1 Steps for submitting the NIH budget

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.

13.4 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 equipment's 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 [9].

13.4.1 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 [9].

"Facilities" is defined as depreciation and use allowances, interest on debt associated with certain buildings, equipment and capital improvements, and operation and maintenance expenses. "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.

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 [9]. 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 subrecipient charges (after the first $25,000) [9].

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.

13.4.2 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.

13.5 Modular Budget

A modular budget format can be submitted if the direct cost is less than $250,000/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.

13.6 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.

13.6.1 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 ($10 \times 0.10 = 1.0$). Other issues to be addressed under this section are salary caps, fringe benefits, and senior/key personnel, which involve 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 [9].

The senior/key personnel who are devoting significant effort to the project should be mentioned. Whereas "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 [9].

13.6.2 Equipment, Travel, and Trainee Costs

Equipment is defined by the NIH as an item of property that has an acquisition cost of $5000 or more (unless the organization has established lower levels), an expected service life of more than 1 year, be stand alone and function independently [9]. 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 $5000, 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.

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.

13.6.3 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.

13.6.3.1 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 $1000 do not have to be itemized.

13.6.3.2 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 [9].

13.6.3.3 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 [9].

13.6.3.4 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 [9]. 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" [9].

13.6.3.5 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.

13.6.3.6 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 [9].

13.6.3.7 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 [9].

13.6.3.8 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 post-doctorate stipend level) [9].

13.6.3.9 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 upfront because if the NIH identifies such an item, they will deduct it from your total award.

13.7 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) [9–11]. When using the detailed budget format in this case, each consortium included must have a separate budget form filled out. In addi-

tion, 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) [9].

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 [9–11].

13.8 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. 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].

13.9 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 [12]. Because of the potential of great monetary benefits, approximately three quarters of funding for clinical drug trials in the United States is sponsored by industry rather than the NIH [13]. Industry-sponsored research thus represents the key supply of funding for an increasing number of clinical investigators [12–14].

13.10 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 [15]. Keeping this in mind, we try to explain how best to plan such a budget.

13.11 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 13.1 and 13.2). 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.

In most research institutions, indirect costs are charged as a mandatory, fixed fee which is usually 20–40% of the total direct cost [15].

Table 13.1 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 13.2 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

13.11.1 Institutional Review Board (IRB) Charges

Before embarking on any investigational study in the United States, 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 [15].

13.11.2 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 [15].

13.11.3 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 [15].

13.11.4 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 [15].

13.11.5 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 [15].

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

13.11.6 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 [15].

13.11.7 Patient Follow-Up

Since most clinical trials rely heavily on patient follow-up data, it is 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 [15].

Table 13.3 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

13.12 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 13.3).

13.13 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 upfront to cover the immediate costs of study initiation (i.e., money to be paid to the investigator before the first patient is enrolled) [15].

13.14 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).

13.15 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 upfront whether you are enrolling 10 or 100 patients [15].

13.16 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 [15].

13.17 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 [15]. Thus, it is advisable to have a clause in the contract that addresses such an event.

13.18 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 healthcare 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.

13.19 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, FY 2020 Presidents Budget, NIH website. https://officeofbudget.od.nih.gov/index.htm. Accessed May 2019.
4. Hormos R, Abkowitz JL, Keating A. Facing the NIH funding crisis, how professional societies can help. JAMA. 2012;308(22):2343–4.
5. Moses H III, Dorsey ER. Biomedical research in an age of austerity. JAMA. 2012;308(22): 2341–2.
6. Office of Budget, National Institutes of Health (NIH). Price indexes. HIH website. https://officeofbudget.od.nih.gov/gbiPriceIndexes.html. Accessed 16 May 2019.
7. National institutes of Health. Research portfolio online reporting tools report, funding, research project success rates by type and activity. https://report.nih.gov/success_rates/Success_ByActivity.cfm. Accessed 5 May 2019.
8. NIH Office of Extramural Research, Grants and Funding. NIH Grants Policy statement. NIHWebsite. https://grants.nih.gov/grants/policy/nihgps/HTML5/section_7/7_cost_consideration.htm. Accessed 5 May 2019.
9. US Department of Health and Human Services. Office of Extramural Research, NIH, Grants process Overview, Developing on your Budget. National Institutes of Health Website. https://grants.nih.gov/grants/how-to-apply-application-guide/format-and-write/develop-your-budget.htm. Accessed 13 May 2019.
10. National Institutes of General Medical Sciences. The research funding, STTR program (R41/42). NIH website. https://www.nigms.nih.gov/research/mechanisms/pages/STTR.aspx. Accessed 16 May 2019.
11. National Institutes of General Medical Sciences, Health, Research Funding. SBIR program. National Institutes of Health Website. https://www.nigms.nih.gov/research/mechanisms/pages/SBIR.aspx. Accessed 16 May 2019.
12. Bodenheimer T. Uneasy alliance: clinical investigators and the pharmaceutical industry. N Engl J Med. 2000;342:1539–44.
13. Staropoli JF. funding and practice of biomedical research. JAMA. 2003;290:112.
14. Nathan DG. Clinical research: perceptions, reality, and proposed solutions: National Institutes of Health director's panel on clinical research. JAMA. 1998;280:1427–31.
15. Beal K, Dean J, Chen J. Budget negotiation for industry sponsored clinical trials. Anesth Analg. 2004;99:173–6.

Chapter 14
Regulatory Considerations in Human Subjects Research

H. Richard Alexander Jr. and Howard S. Hochster

14.1 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 the U.S. 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. Richard Alexander Jr. (✉)
Rutgers Cancer Institute of New Jersey, New Brunswick, NJ, USA

Division of Surgical Oncology, Department of Surgery,
Rutgers Robert Wood Johnson Medical School, New Brunswick, NJ, USA
e-mail: Richard.Alexander@Rutgers.edu

H. S. Hochster
Rutgers Cancer Institute of New Jersey, New Brunswick, NJ, USA

Division of Surgical Oncology, Department of Surgery,
Rutgers Robert Wood Johnson Medical School, New Brunswick, NJ, USA

Division of Medical Oncology, Department of Medicine,
Rutgers Robert Wood Johnson Medical School, New Brunswick, NJ, USA

T. M. Pawlik, J. A. Sosa (eds.), *Clinical Trials*, Success in Academic Surgery,
https://doi.org/10.1007/978-3-030-35488-6_14

14.2 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 performed by the Scottish physician, James Lind, aboard the HMS Salisbury in 1747. At that time, scurvy, a condition that is caused by vitamin C deficiency, was common in sailors who were at sea for extended periods 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. Presumably, none of these sailors, enlisted in the Royal Navy, gave informed consent (verbally or in writing). 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 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 United States 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 U.S. government, meals, and free burial insurance. Investigators never told the study participants that they had syphilis. Perhaps the most egregious breach of ethical conduct is the fact that researchers knowingly failed to treat study participants for their syphilis even after the validation in the early 1940s that penicillin could effectively cure this condition. And, despite penicillin becoming 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 others health facilities in the area.

In another egregious example, from 1946 to 1948, the U.S. Public Health Service in collaboration with Guatemalan health authorities conducted a study in which prisoners, soldiers, and mentally ill patients in that country were infected deliberately 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/Guatemala Syphilis Experiment, accessed 25 Mar 2019).

14.2.1 The Nuremberg Code and Helsinki Declaration

Following World War II, the Nuremberg Military Tribunals were held. They were a series of 12 U.S. military tribunals to prosecute war crimes against members of the leadership of Nazi Germany that were perpetrated against prisoners of war including brutal medical experimentation. The tribunals were held in the Nuremberg Palace of Justice, from 1946 to 1949, following the Allied victory in World War II. 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 toward experiment participants. The ten points of the Nuremberg code are highlighted in Table 14.1 (history.nih.gov/research/downloads/nuremberg.pdf, accessed 27 Dec 2012).

Subsequently, in the 1960s, and based on the Nuremberg Code, 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. While the Helsinki Declaration is not a legally binding instrument, it has been used as the basis for legal statues and regulations overseeing human subject 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 was no broadly established set of guiding 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.

14.2.2 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. The purpose of the Commission was to identify the basic ethical principles that 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

Table 14.1 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

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 that are described in Table 14.2 (http://ohsr.od.nih.gov/guidelines/belmont.html).

Table 14.2 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, non-exploitative, and well-considered procedures are administered fairly—the fair distribution of costs and benefits to *potential* research participants—and equally

14.3 Regulation of Human Subjects Research

14.3.1 The Code of Federal Regulations (CFR) and Federal Wide 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 U.S. 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 subject research. The Division of Policy and Assurances administers the Federal Wide 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 U.S. federal department or agency that has adopted the Common Rule. The individual 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 Protections 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 state-

ment 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 U.S. federal department or agency conducting or supporting the research and OHRP.

In January 2019, the revisions adopted in 2017 and finalized in June 2018 (therefore referred to as 2018 revisions) to the "Common Rule" went into effect after more than 6 years of discussion and comment. These were the first changes since 1991. The revisions to the Common Rule allow for more streamlined consent forms in easier and less technical language with the goals of research at the top, posting of consent forms to national websites, information regarding when results would be released and patient notification. In addition, exemptions for continued long-term follow-up of patients on trials otherwise complete considered "minimal risk" will no longer require IRB oversight for data retrieval and analysis. Biospecimen collection and research on de-identified data in warehouses should be easier. Finally, though delayed until 2020, one IRB of record will be responsible for oversight of multi-centered trials at all sites.

14.3.2 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 research to the OHRP or any U.S. 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 clause 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. The IRB must be notified as soon as possible, and such deviations should be approved in advance if time permits.

14.3.3 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 31 Dec 2012). Safeguards to ensure confidentiality of PHI that is collected as part of a research activity are required by the institutional HRPO that 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 14.3).

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 for scientific integrity 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.

Table 14.3 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

14.3.4 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 & 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 U.S. 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 drug is an investigational agent.
- 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. Besides the obvious medical devices, IDEs are also applied to laboratory testing, new assays for drug targets and multi-assay panels. 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, re-label, 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 14.4). The FDA has oversight of clinical 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 GCP 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 U.S., 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

Table 14.4 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, 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)

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 GCP will increase enthusiasm for the credibility of the findings, and in the case of a FDA audit, avoid a publicly available citation for failure to follow GCP guidelines.

14.3.5 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, etc.). Frequently, these materials may be derived from an academic laboratory. In order to comply with GCP, 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.

14.3.6 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 25 Apr 2019).

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/or processes, and the investigator has to comply with all of them.

14.3.7 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 never acceptable; a translator may not translate consent orally for a participant; and the next of kin or power of attorney may not give consent. In most cases, a general "short-form" approved by the IRB in the language of the subject is used to document the consent process, while a translator should be used to translate the actual informed consent form and document the patient's understanding and agreement to participate. 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 or the investigator cited in a national database. Persistent violations may result in investigator disbarment by the FDA.

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, specific toxicity management, 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 importance, and an auditor will

cite the investigator for the deviation. Additionally, deviations must be reported to the IRB according to the individual IRB requirements. 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 anti-emetics should be avoided due to interactions with a drug used in the study, the investigator should prohibit them, but if any anti-emetic will do then specifying a particular medication 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 should determine in advance what constitutes 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. Some kinds of such operational research will be easier to conduct without IRB oversight under the revised Common Rule.

14.4 Conclusions

Regulatory requirements in human subject research have evolved from guiding principles designed to protect the rights and welfare of the research subject. However, they also serve to protect the institution, 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 curing illness and maintain confidence of the public in the value of such research.

Further Reading

Brown SR. Scurvy: how a surgeon, a mariner, and a gentleman solved the greatest medical mystery of the age of sail. New York, NY: St. Martin's Press; 2003. ISBN 0-312-31391-8.

Goodyear MDE, Krleza-Jeric K, Lemmens T. The Declaration of Helsinki. BMJ. 2007;335(7621):624–5. https://doi.org/10.1136/bmj.39339.610000.BE.

Jones J. Bad blood: the Tuskegee syphilis experiment. New York, NY: Free Press; 1981. ISBN:0-02-916676-4.

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.
www.citiprogram.org.
www.fda.gov.
www.wikipedia.org/wiki/Syphilis_experiments_in_Guatemala.

Chapter 15
Publishing Your Clinical Trial

Warren Gasper and Michael Conte

Publication of a clinical trial is the culmination of an enormous effort; one that deserves an accurate and transparent report.

Although the trial protocol and statistical analysis plan serve as the framework for a clear presentation of the results, it is an arduous task to condense the thousands of design details, data points and analyses into a concise manuscript. Appointing a publication committee to outline the manuscript, settle questions of authorship, and select a target journal assures timely publication of the final report. Guidelines, including the CONSORT Statement, assist in writing a report that is complete and provides maximum value to the medical community. After the initial publication, a thoughtful secondary analysis plan not only maximizes knowledge gained from trial data but it can also help mitigate bias when there are unexpected results. Developing a data sharing plan ensures that the trial data continues to be useful long after the primary and secondary analyses are completed.

15.1 Publication Committee

Establishing a publication committee ensures that clinical trial data will be disseminated widely. The committee's task begins with a publication plan for a timely initial publication and ends with a data handling and sharing plan to facilitate analyses past the end of the trial. Furthermore, by adjudicating analysis proposals, vetting authorship contributions, and synthesizing a data handling plan, the committee facilitates oversight of the publication process by the principal investigator.

W. Gasper (✉) · M. Conte
Division of Vascular and Endovascular Surgery, University of California San Francisco, San Francisco, CA, USA
e-mail: warren.gasper@ucsf.edu; michael.conte2@ucsf.edu

T. M. Pawlik, J. A. Sosa (eds.), *Clinical Trials*, Success in Academic Surgery,
https://doi.org/10.1007/978-3-030-35488-6_15

Once the study protocol and statistical analysis plan are completed, the publication committee outlines the tables and figures necessary for reporting the trial results. The manuscript outline should also be based on the best practice reporting guidelines including the recommendations of the International Committee of Medical Journal Editors (ICMJE) [1], Consolidated Standards of Reporting Trials (CONSORT) 2010 Standards [2] and author instructions from individual medical journals. Having an outline of the final manuscript before subject enrollment has two advantages. First, it improves the efficiency of the final manuscript writing process. Second, it verifies that the data collected in the trial will be appropriate for the final analysis and meet the standards for publication. Notably, when reviewing a manuscript for publication, most journals and the ICMJE require confirmation that the study protocol including a statistical analysis plan was finalized and public registration of the trial with an appropriate trial registry, such as ClinicalTrials.gov, occurred *before enrollment of the first subject* [1].

Public disclosure of the study design is critical to establish the trust of clinicians and patients. It is recommended that the publication committee publish an in-depth description of the clinical problem, study population, and protocol before subject enrollment begins. The chapters of this book provide a comprehensive description of how to design a successful clinical trial, and a discussion of protocol design is beyond the scope of this chapter. However, guidelines such as the 33-item checklist in the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) 2013 Statement focus on protocol completeness and are highly recommended [3, 4]. An announcement report provides clinicians and potential subjects the opportunity to carefully consider the study and build support for enrollment. As the clinical trial proceeds, it may also be necessary to publish interval "update" reports, particularly if there are substantive changes to the study design or protocol.

It is inevitable that a clinical trial will lead to proposals for secondary and subanalyses. These studies may represent an additional analysis of the primary data or, in large trials, may be formal sub-studies with independent research protocols that are extensions of the primary trial. Regardless of whether a secondary analysis is specified in the study protocol or is a post-hoc exploratory analysis, the integrity of the primary trial outcomes must be preserved and overlapping publications of the same data avoided. Secondary analysis proposals should describe the relationship to the primary outcomes, indicate the statistical adjustments for multiple analyses of the primary data, and justify the need for a separate publication. The publication committee is responsible for adjudicating proposals for secondary publications, establishing the publication priorities and drafting abstracts or project outlines. Sponsors of clinical trials budget for the publication of the primary trial results and significant secondary analyses, but may not provide financial support for an unending stream of publications. Co-authors who may be interested in a secondary analysis can be assigned a publication and secure the funding necessary for completing it.

Finally, all clinical trials need to develop a data handling and sharing plan. Eventually, study investigators will exhaust the list of secondary analyses and open data sharing of deidentified data ensures that the data will have lasting benefit to the medical community. Decisions about data formatting, storage location, what data

elements will be shared and with whom are best managed by the steering and publication committees at the outset of the trial. Although industry trial sponsors may claim that the data contain proprietary information and push for sequestration of the data, every effort should be made to negotiate for data sharing. Regardless of what the publication and steering committee decides, the ICJME requires the details of a data sharing plan with a clinical trial manuscript submission [1].

15.2 Manuscript Preparation

As the final analysis is completed, the protocol, statistical analysis plan, and outline of figures and tables provide a framework for the expedient completion of the manuscript. For many investigators, achieving the proposed study end point defines a successful study. However, even successful studies have imperfections, and focusing on a positive or promising outcome while overlooking a contradictory or unexpected result leads to bias in the final report. Guidelines discussed below were developed to ensure accurate reporting of trial results, but it is ultimately incumbent upon the principal investigator and publication committee to honestly report the results and, for complete transparency and reproducibility, to consider data sharing.

Although each journal has specific formatting requirements, the need for accurate and complete clinical trial reporting has led to standards for the substance of a clinical trial manuscript [1, 2]. The CONSORT 2010 Statement is a comprehensive 25-item checklist for reporting clinical trial data with rationales and examples for each topic [5]. One checklist item—a flowchart for depicting the number of subjects enrolled, randomized, treated, and included in the final analysis—has proven so effective that many journals require a "CONSORT flowchart" with all clinical trials. A series of "CONSORT extensions" have augmented the CONSORT statement to describe the best practice for reporting pilot/feasibility [6], non-inferiority [7], pragmatic [8], and cluster [9] trial designs [10]. Other initiatives, such as the Enhancing the Quality and Transparency of Health Research (EQUATOR) Network, offer additional guidelines for improving the quality of all research publications [11, 12].

One conspicuous omission from the CONSORT checklist is a statement of subject protection in accordance with the Declaration of Helsinki [13]. Investigators have an ethical obligation to obtain the appropriate approval from a local, regional, or national institutional review board (IRB) or ethics committee for a clinical trial protocol and obtain informed consent from all subjects. Subject privacy is paramount and no identifying information should be included in the final publication without the explicit consent of the subject. The manuscript should include a statement that confirms IRB approval and subject informed consent were obtained. The topic of adverse events and safety monitoring with a focus on the harm due to a clinical trial intervention has been addressed with a CONSORT extension [14].

To ensure that final results accurately reflect the original protocol and analysis plan, journal editors typically request a copy of the trial protocol and statistical plan. If the manuscript is accepted, the protocol may be published as a component of the

supplemental materials. As discussed above, the SPIRIT 2013 Statement checklist is a useful, highly recommended tool for demonstrating the completeness of a study protocol [3]. Similar guidelines for the uniform reporting of statistical analysis plans (SAP) have been developed, including a checklist of recommended items [15]. At a minimum, most journals require a description of the sample size calculations, primary outcome analysis plan with an intention-to-treat analysis, method for handling missing data, and adjustments for multiple testing with secondary analyses.

As important as the research protocol and statistical analysis plan are for interpreting trial results, without a thorough description of the trial intervention, it may be impossible to replicate the study results. Developed as an expansion of items from the CONSORT and SPIRIT Statements, the 12-item Template for Intervention Description and Replication (TIDieR) checklist focuses on details that affect the efficacy and reproducibility of a study intervention [16]. Including details essential for intervention replication, such as research staff training, exact device type, and intervention personalization or modifications, may exceed the word limit for a primary publication and are often appropriate for supplemental materials. Similarly, the Standards for Reporting Diagnostic Accuracy (STARD) 2015 guidelines provide a checklist for complete reporting of diagnostic test accuracy [17, 18]. Finally, authors are encouraged to use the checklist developed for the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) Statement when reporting the results of prognostic or diagnostic prediction models [19].

15.3 Journal Selection

Successful clinical trial results often deserve publication in high impact journals. Disappointing results may not command the same caliber of journal, but publishing the results of a well-designed and performed negative clinical trial is important to the medical community nonetheless. The publication committee should develop a list of target journals and, as the final data collection and analysis are completed, decide where the trial will be published. If there is a plan to present the results of the trial at a meeting, coordination with the journal editors is essential to comply with embargo policies.

15.4 Authorship

All authors must contribute significantly to a vital component of the study (design, data collection, and/or analysis) and contribute to the completion of the manuscript. Individuals who contribute to the execution of the study but not the manuscript are best listed in the Acknowledgements section. The publication committee serves as the arbiter of authorship and decides whether all site investigators are listed as

authors or whether the principle investigator and members of the steering committee are listed as the authors publishing "on behalf of the investigators of X clinical trial." In the latter case, all site investigators are listed in the acknowledgements or supplemental data sections. For secondary publications, the publication committee facilitates decisions about authorship by weighing the merits of a secondary analysis proposal against the actual contribution of an investigator to the study. Regardless of the role, all authors and acknowledged contributors must disclose any conflicts of interest in the final publication.

References

1. International Committee of Medical Journal Editors. Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals. Vancouver, BC: International Committee of Medical Journal Editors; 2018. http://icmje.org/icmje-recommendations.pdf. Accessed 7 May 2019.
2. Schulz KF, Altman DG, Moher D, Group C. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ. 2010;340:c332. https://doi.org/10.1136/bmj.c332.
3. Chan AW, Tetzlaff JM, Altman DG, Laupacis A, Gotzsche PC, Krleza-Jeric K, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med. 2013;158(3):200–7. https://doi.org/10.7326/0003-4819-158-3-201302050-00583.
4. Chan AW, Tetzlaff JM, Gotzsche PC, Altman DG, Mann H, Berlin JA, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. BMJ. 2013;346:e7586. https://doi.org/10.1136/bmj.e7586.
5. Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ. 2010;340:c869. https://doi.org/10.1136/bmj.c869.
6. Eldridge SM, Chan CL, Campbell MJ, Bond CM, Hopewell S, Thabane L, et al. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. BMJ. 2016;355:i5239. https://doi.org/10.1136/bmj.i5239.
7. Piaggio G, Elbourne DR, Pocock SJ, Evans SJ, Altman DG, Group C. Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. JAMA. 2012;308(24):2594–604. https://doi.org/10.1001/jama.2012.87802.
8. Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. BMJ. 2008;337:a2390. https://doi.org/10.1136/bmj.a2390.
9. Campbell MK, Piaggio G, Elbourne DR, Altman DG, Group C. Consort 2010 statement: extension to cluster randomised trials. BMJ. 2012;345:e5661. https://doi.org/10.1136/bmj.e5661.
10. Extensions of the CONSORT Statement. CONSORT website. 2019. http://www.consort-statement.org/extensions. Accessed 7 May 2019.
11. Simera I, Moher D, Hoey J, Schulz KF, Altman DG. The EQUATOR Network and reporting guidelines: helping to achieve high standards in reporting health research studies. Maturitas. 2009;63(1):4–6. https://doi.org/10.1016/j.maturitas.2009.03.011.
12. The EQUATOR Network. Enhancing the quality and transparency of health research. EQUATOR Network website. 2019. http://www.equator-network.org/. Accessed 7 May 2019.
13. World Medical A. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA. 2013;310(20):2191–4. https://doi.org/10.1001/jama.2013.281053.

14. Ioannidis JP, Evans SJ, Gotzsche PC, O'Neill RT, Altman DG, Schulz K, et al. Better reporting of harms in randomized trials: an extension of the CONSORT statement. Ann Intern Med. 2004;141(10):781–8. https://doi.org/10.7326/0003-4819-141-10-200411160-00009.
15. Gamble C, Krishan A, Stocken D, Lewis S, Juszczak E, Dore C, et al. Guidelines for the content of statistical analysis plans in clinical trials. JAMA. 2017;318(23):2337–43. https://doi.org/10.1001/jama.2017.18556.
16. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ. 2014;348:g1687. https://doi.org/10.1136/bmj.g1687.
17. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ. 2015;351:h5527. https://doi.org/10.1136/bmj.h5527.
18. Cohen JF, Korevaar DA, Altman DG, Bruns DE, Gatsonis CA, Hooft L, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. BMJ Open. 2016;6(11):e012799. https://doi.org/10.1136/bmjopen-2016-012799.
19. Collins GS, Reitsma JB, Altman DG, Moons KG, Group T. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. The TRIPOD Group. Circulation. 2015;131(2):211–9. https://doi.org/10.1161/CIRCULATIONAHA.114.014508.

Chapter 16
Pragmatic Clinical Trials

Peter C. Minneci and Katherine J. Deans

16.1 What Is a Pragmatic Clinical Trial?

Pragmatic clinical trials represent an extension of the clinical research continuum to produce evidence that supports clinical decision making for individual patients. In general, clinical trials are used to investigate the safety, efficacy, and/or effectiveness of treatments. The goal of a pragmatic clinical trial is to determine the effectiveness of treatment strategies in clinical practice.

According to the NIH, a clinical trial prospectively assigns human subjects to one or more interventions and evaluates their effects on health-related outcomes [1]. Clinical trials can be classified in several ways based on their purpose, phase, or design. The NIH organizes clinical trials into five categories based on their purpose: (a) prevention trials, (b) screening trials, (c) diagnostic trials, (d) treatment trials, and (e) trials [2]. Although surgeons may perform clinical trials in any of these categories, surgical trials are the most common treatment trials. These types of trials test a pre-specified hypothesis about potential treatment differences between two or more treatments by enrolling participants from a specified patient population [2].

Treatment trials evaluating new treatments are often conducted in phases [2]. Each phase has a different purpose with subsequent phases using the results from previous phases to inform the trial. Phase I trials test a potential treatment in humans for the first time to establish safety including a safe dose range and identifying side effects. Phase II trials begin to establish efficacy and further evaluate a treatment's

P. C. Minneci · K. J. Deans (✉)
Center for Surgical Outcomes Research, The Research Institute and Department of Surgery, Nationwide Children's Hospital, Columbus, OH, USA

Department of Surgery, The Ohio State University College of Medicine, Columbus, OH, USA
e-mail: Peter.minneci@nationwidechildrens.org; Katherine.deans@nationwidechildrens.org

T. M. Pawlik, J. A. Sosa (eds.), *Clinical Trials*, Success in Academic Surgery,
https://doi.org/10.1007/978-3-030-35488-6_16

safety. Phase III trials are typically large randomized controlled trials that establish a treatment's efficacy compared to placebo or the current standard treatment. Phase IV trials are post-market trials to evaluate long-term side effects and establish effectiveness in clinical practice. Pragmatic clinical trials are large trials that would be categorized as either phase III or IV trials. Their goal is to provide evidence of the effectiveness of a treatment in routine clinical practice.

The driving force for the development of more pragmatic clinical trials is the fact that many trials do not adequately inform clinical practice because they were designed to establish efficacy and not effectiveness [3, 4]. The typical RCT is performed by experienced researchers in highly selected patient populations under highly controlled conditions. This may lead to an overestimation of the beneficial effects of a treatment and possibly an underestimation of is harmful effects [3, 4]. Furthermore, results obtained from these trials may not be applicable to "real-world" practice because most trials are performed under artificial conditions with study populations that are not reflective of the broader population of patients with the disease [5].

16.2 When Should a Pragmatic Trial Be Considered? Efficacy Versus Effectiveness RCTs

The "gold" standard of evidence-based medicine is the randomized controlled trial (RCT) which directly compares a treatment to a "control." A RCT can be broadly categorized as either an efficacy or effectiveness trial. Efficacy trials, also referred to as "explanatory" trials, are designed to test causal research hypotheses [6]. They are designed to determine the effects of a treatment under ideal circumstances in relatively homogeneous patient populations. In contrast, effectiveness trials, also referred to as pragmatic trials, aim to generate results that can be used by clinicians to choose between treatments [6]. They are designed to determine the effects of a treatment in clinical care under the usual conditions in which it is administered.

Randomization in a clinical trial can control for selection bias and allow a causal link to be established between a treatment and changes in the primary outcome [6, 7]. RCTs have high internal validity, but generalizability of the results of a RCT can be limited depending on the restrictiveness of the treatment protocol or the included patient population [7]. Efficacy or explanatory trials designed to establish a causal link between a treatment and outcome may be too restrictive to generate results that can be used broadly in clinical practice. Trials using narrow inclusion and exclusion criteria may limit the included patient population to only a small subset of those patients treated in clinical practice. Therefore, the benefit of a treatment in subsets of patients excluded from the trial or in the broader patient population remains unknown.

16.3 Why Do We Need Pragmatic Clinical Trials?

Pragmatic clinical trials are needed because results from traditional randomized controlled trials oftentimes do not include patients with complex or comorbid conditions who might benefit the most from the study treatment [7]. In addition, traditional randomized controlled trials are rarely performed in typical clinical settings and oftentimes ask questions and assess outcomes that are not important to patients, clinicians, or policy makers. Furthermore, traditional RCTS are expensive, take years to complete, and the results can take many years to be disseminated and implemented [7]. These issues make the results of traditional RCTs difficult to translate into everyday clinical practice and often times lead to a dramatic decrease in the effectiveness of a treatment when it used in clinical practice.

In contrast, pragmatic clinical trials are practical, inclusive, engaged, and relevant [7]. They are designed with an emphasis on using a treatment in clinical practice with a goal of broader implementation. They typically test a treatment strategy that can be implemented into every day clinical practice in a diverse patient population. Relevant stakeholders, including patients, clinicians, and healthcare systems, are involved in designing the study, interpreting the results, and implementing the findings [7]. Therefore, the questions they investigate tend to be more patient centric.

16.4 The Continuum of Clinical Trials

Explanatory trials study how well treatments work in well-specified and typically ideal conditions; whereas, pragmatic trials study how well treatments work in usual clinical settings. However, no trial is completely pragmatic or explanatory. Rather, clinical trials fall along a continuum between the traditional explanatory efficacy trial to a purely pragmatic effectiveness trial (Fig. 16.1) [7, 8]. Important differences between explanatory and pragmatic trials are shown in Table 16.1 [7].

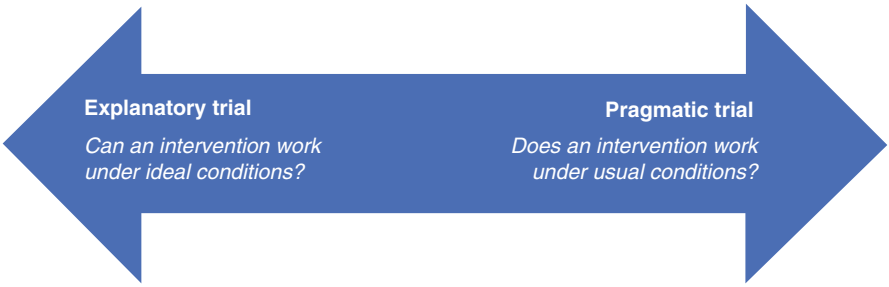


Fig. 16.1 Continuum between explanatory and pragmatic trials (Reprinted with permission from the NIH Health Care Systems Research Collaboratory [7])

Table 16.1 Key differences between randomized controlled trials (RCT) and pragmatic clinical trials (PCT)

	Explanatory trial	Pragmatic trial
Overview	A traditional RCT tests a hypothesis under ideal conditions	A PCT compares treatments under everyday clinical conditions
Goals	To determine causes and effects of treatment	To improve practice and inform clinical and policy decisions
Design	Tests the intervention against placebo using rigid study protocols and minimal variation	Tests two or more real-world treatments using flexible protocols and local customization
Participants	Highly defined and carefully selected	More representative because eligibility criteria are less strict
Measures	Require data collection outside routine clinical care	Brief and designed so data can be easily collected in clinical settings
Results	Rarely relevant to everyday practice	Useful in everyday practice, especially clinical decision making

Reprinted with permission from the NIH Health Care Systems Research Collaboratory [7]

Explanatory trials attempt to establish efficacy under ideal circumstances. Their goal is to assess the causal relationship of the treatment and outcome. They have rigid study protocols to limit variation and have selective inclusion criteria to ensure a homogeneous patient population. Research data collection is usually in addition to usual care, and the outcomes assessed are based on answering the specific research question. In contrast, pragmatic trials attempt to establish the relative effectiveness of two treatment strategies in usual care. Their goal is to generate results that can be used to make clinical decisions. In order to maximize the generalizability of the results, their study protocols should be reflective of usual care and they should have broad inclusion criteria to allow for a study population representative of the entire population with the disease. Data collection is designed to be as much a part of usual care as possible, and the outcomes assessed are meant to be clinically relevant to the patients, clinicians, and policy makers.

16.5 Core Characteristics of Pragmatic Trials

As first described by Schwartz and Lellouch, pragmatic trials attempt to establish the relative effectiveness of treatments in real- life clinical practice [3, 4]. Core characteristics of pragmatic trials include: (1) attempting to answer questions that are from, and important to, stakeholders, (2) assessing multiple outcomes that are important to shared decision-making and policymakers, (3) comparing the treatments to real-world alternatives, and (4) being performed in diverse representative populations and in multiple heterogeneous settings [7].

Pragmatism can be incorporated into most elements of trial design. Table 16.2 details the nine dimensions for assessing the level of pragmatism in a clinical trial as proposed in the pragmatic explanatory continuum indicator summary 2

Table 16.2 PRECIS-2 domains reprinted with permission from the NIH Health Care Systems Research Collaboratory [9, 10]

PRECIS-2 domain	Description
Eligibility criteria	Who is selected to participate in the trial? A pragmatic approach would be to include anyone with the condition of interest who is likely to be a candidate for the intervention if it were being provided in usual care for this condition
Recruitment	How are participants recruited into the trial? A pragmatic approach for patient recruitment would be through usual appointments at a diverse range of clinics to increase the applicability of the trial results. (Note that with PCTs, trial participants can also be groups of care providers or health systems)
Setting	Where is the trial being done? Several characteristics of the setting could affect the applicability of the results, including geography, healthcare system, country, and the socioeconomic and ethnic mix of the population. A pragmatic approach would be to do the trial in an identical setting to which you intend the results to be applied
Organization	What expertise and resources are needed to deliver the intervention? A more pragmatic design would incorporate the intervention into the usual organization of care (e.g., clinical workflow) for the condition of interest, making use of no more than the existing healthcare staff and resources in that setting
Flexibility (delivery)	How should the intervention be delivered? The most pragmatic approach to deliver flexibility would leave the details of how to implement the intervention up to providers, as happens in usual care. Thus, the methodology of how to deliver an intervention is not rigidly prescriptive in the protocol
Flexibility (adherence)	What measures are in place to ensure participants adhere to the intervention? A pragmatic approach would allow for full flexibility in how end user recipients engage with the intervention
Follow-up	How closely are participants followed up? A pragmatic design would be to have no more follow-up of recipients than would be the case in usual care. Most pragmatic would be to obtain outcome data by other means such as the EHR or other usual data to measure mortality or hospital admissions
Primary outcome	How relevant is it to participants? A pragmatic approach would be to select an outcome that is of obvious importance from the perspectives of all stakeholders. For example, "an intervention that aims to reduce falls in elderly people living independently in the community should have as its primary outcome the number of falls in the elderly living independently in the community. This outcome has meaning to patients, their relatives and friends, healthcare professionals, and policymakers" [9]
Primary analysis	To what extent are all data included? A pragmatic approach to analysis would be to make no special allowance for non-adherence, practice variability, etc. In other words, the pragmatic approach to the analysis would typically be an intention-to-treat analysis using all available data

(PRECIS-2) tool [9, 10]. A pragmatic clinical trial attempts to design each aspect of the trial to be reflective of who, how, and where the treatment would be used in clinical practice and minimize differences between research and clinical practice. Eligibility criteria are designed to be inclusive of patients who would be receiving the treatment in clinical practice. Recruitment is ideally performed in a diverse group of both community and academic clinics or hospitals to improve the generalizability of the results. Performance of the trial, including administration of the treatment and data collection, would be incorporated into existing clinical workflow. Furthermore, monitoring of treatment adherence is flexible. Treatment modifications are at the discretion of the clinicians as opposed to rigid study-related protocols. Follow-up and outcome assessment should be part of usual care and recorded as part of encounters in the electronic health record. The primary outcome for a pragmatic trial should be important to patients and all other relevant stakeholders. The primary analysis should be intention-to-treat using all available data.

16.6 Unique Characteristics of Pragmatic Clinical Trials

Performance of a pragmatic trial often works best when the treatment is assigned at a group level rather than at the patient level. For this reason, cluster randomization is often used [3]. In cluster randomization, groups of patients treated within a single clinic or hospital are randomly assigned to the same treatment. The different clinics or hospitals participating in the trial are randomized to be either intervention or control treatment centers. All patients enrolled in the trial receive the assigned treatment based on center randomization. Outcomes can then be assessed at the individual patient level (cluster–individual trial) or at the cluster level (cluster–cluster trial) [3, 11, 12]. Cluster–cluster trials allow for the possibility of a complete waiver of consent for enrollment in the trial, whereas cluster–individual trials allow for the option of waiving consent for the treatment but attaining consent for follow-up. Another trial design option used in pragmatic trials is a stepped-wedge cluster design [3, 13]. This is a type of pre-post study design in which all sites will transition from the control treatment to the intervention treatment, but the timing and order of when sites transition is variable and randomly assigned. This cluster design can be used with the option to waive individual consent. With either of these types of cluster trials, inclusion of more diverse and representative clinics or hospitals will increase the generalizability of the results.

In addition to using cluster randomization, another common feature of pragmatic trials is the use of the electronic health record to improve efficiency and minimize cost. The electronic health record can be used to identify patients for recruitment, perform baseline and outcome data collection, monitor participants, and communicate with participants to perform follow-up [7]. The development of Learning Health Systems can support and promote the performance of pragmatic clinical

trials. A Learning Health System allows for data from the electronic health record to be collected for a trial while allowing each participating institution to maintain control over their individual data [10, 14]. The potential for performing pragmatic clinical trials using Learning Health Systems has been supported by the Patient-Centered Outcomes Research Institute (PCORI). PCORI funded the development of PCORnet, a National Patient-Centered Clinical Research Network, which links integrated health systems and "patient-powered" networks. The purpose of PCORnet is to facilitate multi-site, observational, and interventional comparative effectiveness research across linked clinical data research networks and patient-powered research networks [10, 14].

16.7 Challenges in Pragmatic Trials

Several challenges can be encountered when designing a pragmatic trial. Pragmatic trials seek to include patient populations similar to those cases where the treatment will be used in clinical practice; however, this may not be known for new treatments [3]. Also, pragmatic trials attempt to include clinics and hospitals that reflect the heterogeneity of clinical practice. However, all clinical trials need local investigators who are engaged and take responsibility for performing the trial. Many clinicians practicing in nonacademic settings may not have familiarity or resources to perform the additional responsibilities associated with performing a clinical trial [3]. This can lead to underrepresentation of community sites or poor compliance with research protocols at these sites. Heterogeneity of sites may also be especially harmful in surgical trials [3]. If a trial is investigating a complex surgical procedure which is more commonly performed at high volume centers (e.g., a Whipple procedure for pancreatic cancer), then including lower volume sites may limit the usefulness of the results. Finally, many pragmatic trials are not blinded. This allows for potential bias during outcome assessment. Strategies to minimize the potential for bias include using major events, such as mortality, as the primary outcome or having the assessors of outcomes blinded to treatment assignments [3].

16.8 Conclusions

Pragmatic clinical trials attempt to produce evidence to support clinical decision making for individual patients. As compared to exploratory trials, pragmatic trials investigate the effectiveness of a practical treatment strategy in a more diverse patient population and assess outcomes that can inform patients and clinicians trying to make a treatment decision in clinical practice.

References

1. National Institutes of Health. NIH's definition of a clinical trial 2017. 2017. https://grants.nih.gov/policy/clinical-trials/definition.htm. Accessed 16 Sep 2018.
2. National Institutes of Health. NIH clinical research trials and you: the basics 2017. 2017. https://www.nih.gov/health-information/nih-clinical-research-trials-you/basics. Accessed 16 Nov 2018.
3. Ford I, Norrie J. Pragmatic trials. N Engl J Med. 2016;375(5):454–63.
4. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. J Chronic Dis. 1967;20(8):637–48.
5. Ioannidis JP. Why most published research findings are false. PLoS Med. 2005;2(8):e124.
6. Weinfurt K. What is a pragmatic clinical trial: pragmatic elements: an introduction to PRECIS-2. In: Rethinking clinical trials: a living textbook of pragmatic clinical trials. Bethesda, MD: NIH Health Care Systems Research Collaboratory; 2018. https://doi.org/10.28929/092. https://rethinkingclinicaltrials.org/chapters/pragmatic-clinical-trial/post-6366/.
7. National Institutes of Health Care Systems Research Collaboratory. Introduction to pragmatic clinical trials: how pragmatic clinical trials bridge the gap between research and care. Presentation from rethinking clinical trials. A living textbook of pragmatic clinical trials. Bethesda, MD: NIH Health Care Systems Research Collaboratory. https://www.nihcollaboratory.org/Products/Introduction%20to%20pragmatic%20clinical%20trials.pdf.
8. Thorpe KE, Zwarenstein M, Oxman AD, Treweek S, Furberg CD, Altman DG, et al. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. Can Med Assoc J. 2009;180(10):E47–57.
9. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. BMJ. 2015;350:h2147.
10. Weinfurt K. What is a pragmatic clinical trial: definition of a pragmatic clinical trial. In: Rethinking clinical trials: a living textbook of pragmatic clinical trials. Bethesda, MD: NIH Health Care Systems Research Collaboratory; 2018. https://doi.org/10.28929/090. https://rethinkingclinicaltrials.org/chapters/pragmatic-clinical-trial/what-is-a-pragmatic-clinical-trial-2/.
11. Califf RM, Sugarman J. Exploring the ethical and regulatory issues in pragmatic clinical trials. Clin Trials. 2015;12(5):436–41.
12. Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. BMJ. 2008;337:a2390.
13. Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. Contemp Clin Trials. 2007;28(2):182–91.
14. Fleurence RL, Curtis LH, Califf RM, Platt R, Selby JV, Brown JS. Launching PCORnet, a national patient-centered clinical research network. J Am Med Inform Assoc. 2014;21(4):578–82.

Chapter 17
Cooperative Clinical Trials

Casey J. Allen, Giampaolo Perri, and Matthew H. G. Katz

17.1 Overview

In clinical trials, researchers carefully and methodically test drugs, medical devices, screening approaches, behavioral modifications, and other interventions. Because clinical trials conducted at single institutions may be vulnerable to biases and methodologic pitfalls that may result in poorly generalizable or invalid results [1], carefully designed multicenter clinical trials (MCTs, or cooperative trials) are being increasingly performed to advance medical and surgical science. Over 35 years ago, Levin and colleagues provided examples of "the importance and the need for well-designed cooperative efforts to achieve clinical investigations of the highest quality" [2].

Cooperative trials are desirable when a large number of study participants are required to answer a research question, and a single site does not anticipate the ability to enroll a sufficient number [3]. An MCT design may also be employed to accelerate the rate of trial accrual relative to what might be possible by enrolling patients at a single site, to study a rare disease, to increase the generalizability of a study's conclusions by enrolling a more demographically heterogeneous patient population, and to increase the speed with which knowledge is shared. However, some disadvantages to cooperative trials exist. For example, they are typically more complex and more expensive to perform than single-center studies, and the regulatory burden associated with these trials is generally significant.

C. J. Allen · G. Perri · M. H. G. Katz (✉)
Department of Surgical Oncology, University of Texas MD Anderson Cancer Center, Houston, TX, USA
e-mail: mhgkatz@mdanderson.org

T. M. Pawlik, J. A. Sosa (eds.), *Clinical Trials*, Success in Academic Surgery,
https://doi.org/10.1007/978-3-030-35488-6_17

Cooperative trials play a major role in advancing knowledge in the field of surgery over a wide range of diseases and treatment approaches. With regard to cancer research, for example, clinical trials are essential to the development of new methods to prevent, diagnose, and treat disease. We have now established a greater understanding of the molecular and genetic drivers of how cancers develop, grow, and spread. With the ability to quickly and cheaply sequence the genome of an individual patient's tumor, precision medicine has become a reality. We can identify biomarkers to help detect cancers early and guide therapy, and researchers are developing new targeted and immune-based therapies in hopes to improve the results of treatment of previously difficult-to-treat malignancies [4]. Well-designed and executed cooperative trials are increasingly necessary to test the capabilities of these advances. In addition to testing novel interventions, trials may also be conducted to determine the best use of existing interventions (e.g., surgery, radiotherapy, chemotherapy), to test methods to improve end-of-life care, and to assess whether specific treatment approaches can improve patients' quality of life.

In addition to their ability to help define patient care and influence clinical practice, cooperative clinical trials expand the number of clinically oriented investigators who may contribute meaningfully to medical and surgical science. While in the past, clinical trials were most often conducted at major academic centers, community-based clinical practices can successfully participate in cooperative MCTs and, indeed, community-based centers now represent a major source of patient accrual to large clinical trials nationwide.

Although cooperative trials can be performed in a variety of surgical research settings, we place primary focus herein on cooperative cancer trials. We discuss the current organizational structure of the national collaborative network for cancer, describe how MCTs are conducted within this framework, and illustrate steps involved in developing a protocol and seeing its activation. For detailed information on these various complex components, we provide references and recommendation for additional reading [5, 6].

17.2 National Cancer Institute Cooperative Group System

Within the context of cancer research, a "cooperative group" consists of institutions, physicians, and researchers that collaborate on clinical trials which focus on malignant disease. Large cooperative groups have long been supported by the National Cancer Institute (NCI). One of the primary goals of NCI-funded cooperative groups is to promote investigations that are not typically prioritized and supported by industry. In this regard, trials particularly well-suited to conduct within the NCI-funded cooperative groups include, but are not limited to studies of multi-modality treatment programs; studies in which multiple agents from different sponsors are combined into novel regimens; and studies of screening, diagnostic, and prevention strategies.

17.2.1 History of the NCI-Funded Cooperative Groups: From Cooperative Groups to Network Groups

Until 2010, the NCI's Cooperative Group Program was composed of 10 discrete groups: Radiation Therapy Oncology Group (RTOG), Southwest Oncology Group (SWOG), American College of Surgeons Oncology Group (ACOSOG), North Central Cancer Treatment Group (NCCTG), National Surgical Adjuvant Breast and Bowel Project (NSABP), Gynecologic Oncology Group (GOG), American College of Radiology Imaging Network (ACRIN), Children's Oncology Group (COG), Eastern Cooperative Oncology Group (ECOG), and Cancer and Leukemia Group B (CALGB). Each group had its own organizational structure, statistics and data operations, tumor banks, member sites, and disease site committees (Fig. 17.1). The 3,100 institutions and 14,000 investigators who participated in these cooperative groups enrolled more than 25,000 patients in clinical trials each year.

Through its scale and scope, the cooperative group system was responsible for significant progress in oncology care. Practice-changing surgical studies conducted within this system included ACOSOG Z9001 which clearly demonstrated the value of Imatinib following surgery for gastrointestinal stromal tumors [7] and ACOSOG Z0011 which showed that sentinel lymph node dissection is not inferior to axillary dissection in sentinel lymph node-positive breast cancer [8]. In the medical oncology field, successful studies were responsible for FDA approval, for only two

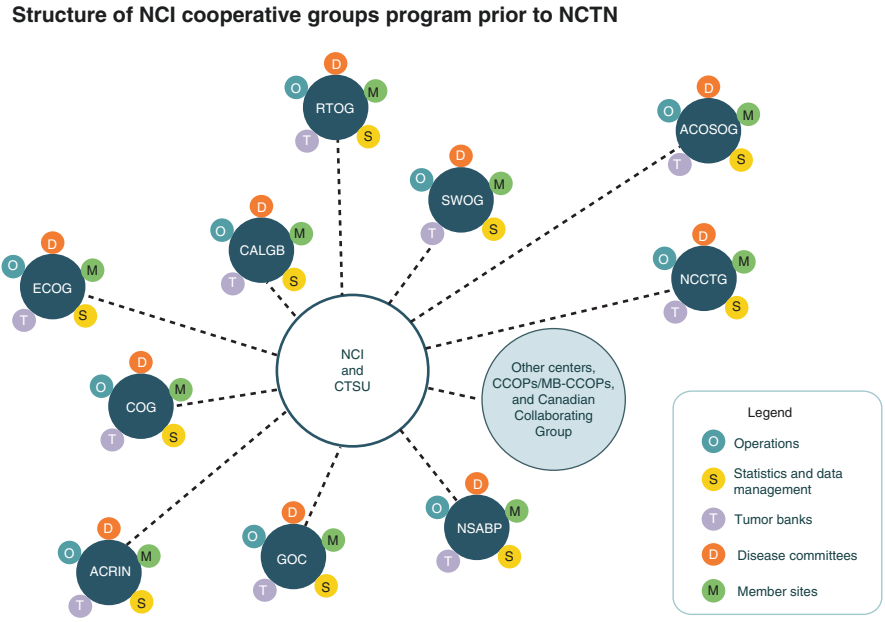


Fig. 17.1 Original NCI cooperative group program structure [24]

examples, of bevacizumab for patients with colon cancer [9] and non-small cell lung cancer [10], and trastuzumab as adjuvant therapy for early-stage Her2+ breast cancer [11]. These successes notwithstanding, several drawbacks were inherent in the cooperative group system's structure. Foremost among these were significant organizational and procedural redundancies which led to significant delays in trial development, initiation, and enrollment. Furthermore, the environment generally favored competition between groups, as opposed to fostering intergroup collaboration and team science.

In 2010, the Institute of Medicine (IOM) provided a report on the state of the system for conducting cancer clinical trials [12]. Stated within the report, "Clinical trials are essential for developing new and improved therapies for patients with cancer. However, the system for conducting cancer clinical trials is approaching a state of crisis. Changes are urgently needed. If the clinical trials system does not improve its efficiency and effectiveness, the introduction of new treatments for cancer will be delayed and patient lives will be lost unnecessarily." The IOM report emphasized the need for a more efficient system that could respond more rapidly to scientific opportunities. The report also emphasized how trials were not keeping up with the overall pace of scientific discovery. So, the IOM provided specific goals as to improve the system for conducting cancer MCTs (Table 17.1). Fundamental to its recommendations was the support of initiatives designed to centralize operations and functions common to all cooperative groups within the system. For example, the IOM recommended consolidating the multiple operations and statistical centers into one central organization. The IOM also mandated efforts to streamline the process of developing trials and implemented mandatory deadlines and other hard stops in the development and activation process.

Guided by the recommendations of the 2010 IOM report, and after consultation and coordination with many stakeholders, the NCI transformed its longstanding Cooperative Group program into the new National Clinical Trials Network (NCTN) and its organizational structure that exists today. The design and implementation of the NCTN incorporated feedback from Cooperative Group investigators, NCI Comprehensive Cancer Center directors, several NCI working groups, leading cancer researchers, industry representatives, and patient advocates [13]. For a detailed history of the structure and the evolution of U.S. Cooperative Group Trials, please refer to additional reading [14, 15].

Table 17.1 2010 Institute of Medicine goals

Goal I: Improve the speed and efficiency of the design, launch, and conduct of clinical trials
Goal II: Incorporate innovative science and trial design into cancer clinical trials
Goal III: Improve prioritization, selection, support, and completion of cancer clinical trials
Goal IV: Incentivize the participation of patients and physician in clinical trials

17.2.2 National Clinical Trials Network

The NCTN includes a network of organizations and clinicians that conduct large phase II and phase III clinical trials, provides infrastructure for MCTs at over 3,000 sites across the United States and Canada, and supports many precision medicine trials [16]. The NCTN structure now includes five U.S. groups and a single Canadian group. The five U.S. network Groups are Alliance for Clinical Trials in Oncology [17], ECOG-ACRIN Cancer Research Group [18], NRG Oncology [19], SWOG [20], and the Children's Oncology Group (COG) [21]. The Canadian Group is the Canadian Cancer Trials Group (CCTG) [22]. Although some competition between its member groups continues to exist, collaboration is now emphasized and has become a priority.

Individual participating institutions can belong to one or more than one cooperative group, and membership in one group allows investigators at any accredited member institution to participate in almost any trial led by any of the NCTN groups. Centralized operations centers are responsible for managing the various subcommittees of each group, as well as providing oversight of the many protocols within the groups.

17.2.3 Lead Academic Participating Sites (LAPS)

Lead Academic Participating Site (LAPS) grants provide direct funding to large academic research institutions that have shown scientific leadership in the design and conduct of clinical trials as well as the ability to enroll high numbers of patients into NCTN trials. LAPS grants support the research staff required to manage patient enrollment and data management. LAPS grants also fund scientific and administrative leadership, as well as education and training costs for staff to better promote patient enrollment. Thirty U.S. academic institutions have been awarded a LAPS grant, and most of the awardees are NCI-Designated Cancer Centers [23]. The current LAPS grantees are listed on the NCI website [24].

17.2.4 NCI Community Oncology Research Program (NCORP)

The NCI provides a robust infrastructure to encourage the participation of community providers and centers in the national network. NCORP is a network of community-based healthcare systems across the United States that conducts MCTs. The sites are consortia of researchers, hospitals, physician practices, academic medical centers, and other groups that provide healthcare services in communities. By

providing access to clinical trials and the benefits of the latest research, the NCORP allows populations that have been historically under-represented in cancer clinical research to gain access to various cutting-edge national trials. These sites can receive independent awards from the NCI [25] or they can receive appropriate reimbursement directly from the network group in which they are affiliated.

17.2.5 Scientific and Administrative Oversight and Additional Components of the NCTN

Whereas the individual cooperative groups develop and conduct each trial, and member sites accrue participating patients, the NCI provides overall administrative and scientific leadership. Logistically, the cooperative groups propose new trial concepts to the NCI Disease/Imaging Steering Committees [26] which then evaluate, prioritize, and eventually approve those trial concepts with the highest potential for scientific impact. These committees also establish and review strategic priorities for a given disease site or research area, maintain a balanced research portfolio across the NCTN, and form task forces and working groups to focus on specific diseases. Membership is composed of representatives from each of the NCTN cooperative groups, NCORP representatives, community oncologists, biostatisticians, patient advocates, and others. The Scientific Steering Committees and associated Task Forces meet monthly, generally by teleconference.

The Coordinating Center for Clinical Trials (CCCT) [27] manages the scientific steering committees and helps to coordinate NCI clinical trials oversight committees including the Clinical Trials and Translational Research Advisory Committee (CTAC). In order to perform trials that focus on using large populations to identify tumor specifics that may respond to novel therapies, the CTAC oversees the organizational structure, funding, and long-term strategic direction of these translational trials [28].

There is immense complexity within the interplay between the various network groups and members and their administrative support, oversight committees, and support services. Figure 17.2 depicts the current NCTN organizational structure.

17.3 NCI-Funded NCTN Cooperative Groups

17.3.1 Group Meetings

In-person group meetings are typically held by each of the cooperative groups twice yearly, in the fall and spring. These meetings are used to establish research goals and review progress toward those goals, to educate the group membership, and to review and disseminate new scientific data. All group members, including physicians, research staff, and other personnel, are encouraged to attend the bian-

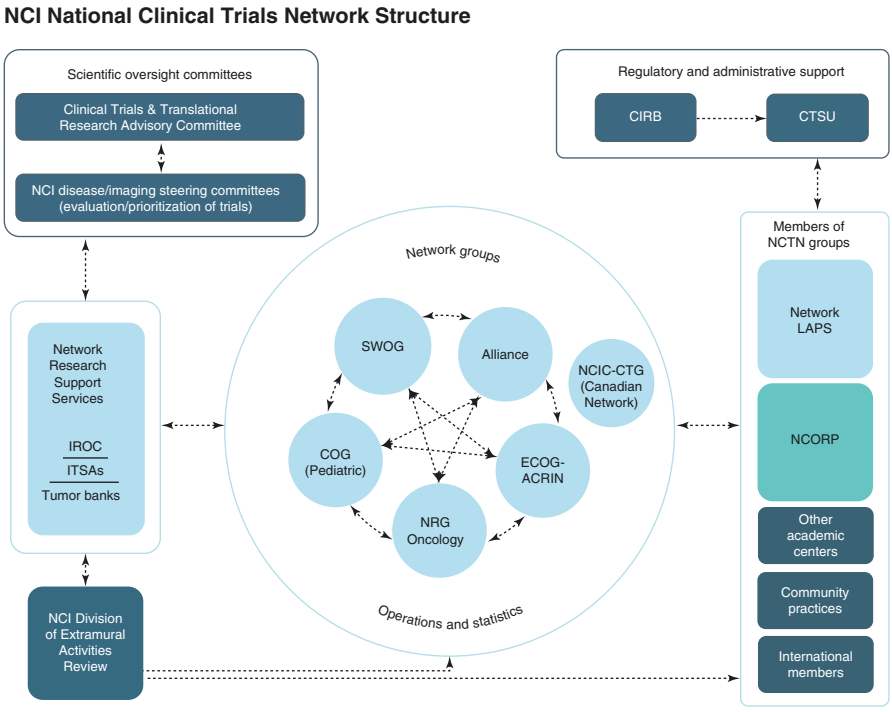


Fig. 17.2 The current NCTN structure [24] *NCORP* NCI Community Oncology Research Program, *LAPS* Lead Academic Participating Sites, *IROC* Imaging and Radiation Oncology Core, *ITSAs* Integrated Translational Science Awards, *CIRB* Central Institutional Review Board, *CTSU* Cancer Trials Support Unit, *CCCT* Coordinating Center for Clinical Trials

nual group meetings. Although specific events and workshops may be only open by invitation, most are open to the entire membership. Travel and registration to the group meetings may be supported by the group's committees or workgroups and is often offered to ranking committee members.

17.3.2 Scientific Committees

The research agenda of each cooperative group is designed and implemented by its scientific committees. These are primarily aligned with a cancer disease site or group of sites (e.g., breast, gastrointestinal, and thoracic) but others include committees focused on disciplines (e.g., early therapeutics, prevention, and epidemiology) or treatment modalities (e.g., radiation oncology). The number and array of committees varies by group. Membership in the scientific committees is multidisciplinary. Each of the scientific committees is led by a committee chair who is primarily responsible for the activities of the committee. The primary functions of the scientific committees include defining scientific programs and developing,

reviewing, and approving new research protocols. Committees often have liaisons to other related committees who facilitate interaction and collaboration.

The actual work of the scientific committees is primarily conducted by smaller subcommittees and workgroups which meet, typically monthly, by teleconference. It is during these conferences that initial ideas for studies are formulated, vetted, and prioritized. Studies with significant merit are brought forward to the scientific group for evaluation at the annual or semi-annual group meetings.

17.3.3 Administrative and Support Committees

Administrative and research support committees, as their name would suggest, support the work of the scientific committees. Such committees generally include an imaging committee, a publication committee, a pharmacy committee, a quality control committee, and others. The specific functions of these committees include providing scientific input, instituting quality control measures, developing educational programs, and providing administrative support. The administrative committees do not produce scientific study protocols on their own.

17.3.4 Surgical Committees

Because the cooperative groups are multidisciplinary groups, the list of each group's scientific committees does not include a committee with a specific focus on surgical trials. However, in some of the cooperative groups (e.g., Southwest Oncology Group), a discrete Surgery Committee exists among the roster of major administrative committees. In others, a subcommittee with a surgical focus may exist within each of the primary disease-site-focused scientific committees.

In the Alliance for Surgical Trials in Oncology, a unique collaboration with the American College of Surgeons also exists: the Alliance/American College of Surgeons Clinical Research Program (ACS-CRP). The mission of this program is to reduce the impact of cancer by increasing knowledge and awareness of new evidence and practice standards in surgery, to increase the participation of community oncology surgeons in cancer research activities, to implement evidence-based practice in surgical oncology, and to create opportunities for meaningful health services research. The ACS-CRP is composed of four committees: Education, Dissemination and Implementation, Cancer Care Standards Development, and Cancer Care Delivery Research. One of the primary products of this program is the Operative Standards for Cancer Surgery manuals, which define technical standards for the conduct of major oncologic operations [29, 30]. The ACS-CRP provides an excellent opportunity for early-career surgeon investigators to become involved with cooperative group surgical research.

17.3.5 Other Committees and Opportunities for the Surgeon

Other opportunities for the surgeon exist within each of the cooperative groups. The Alliance for Clinical Trials in Oncology, for example, has major programs focused on other aspects of clinical cancer research applicable to surgeons. These include the Translational Research Program, which focuses on molecularly driven oncology and focuses on the integration of translational endpoints into clinical trials. It also coordinates the Alliance Biorepository System and provides oversight to translational correlative studies. Similar programs exist within each of the other cooperative groups.

17.4 Protocol Development Within the NCTN: Idea to Activation

The following series of steps represent the general approach to the planning and conduct of a treatment trial within an NCTN cooperative group, from developing an idea through activation (Fig. 17.3). Other types of trials, or treatment trials developed within other multicenter settings, may use a similar general algorithm.

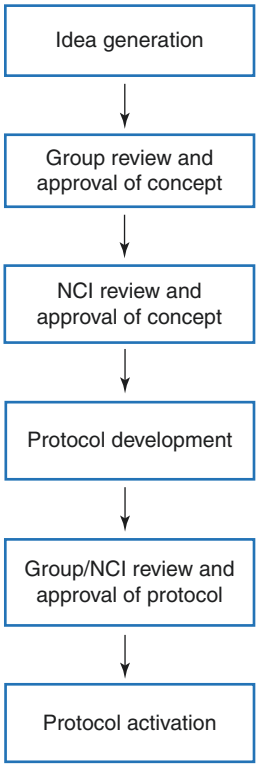


Fig. 17.3 General process used to develop a clinical trial within the NCTN

17.4.1 Idea Generation and Concept Development

The first methodologic step is the formulation of a study question. Each clinical trial should be designed to answer to a single primary clinical question that must be carefully selected and clearly defined. A clearly stated question encourages proper trial design and determines the credibility of any findings; alternatively, a poorly conceived study question will likely destroy a trial before it even begins.

Protocols designed to answer a specific study question are produced within the scientific committees of each cooperative group. Once a clinical question is proposed by a committee member, the member brings the idea to the committee chair for support. Colleagues and collaborators on the sponsoring committee are assembled, typically within a subcommittee or working group, to refine the clinical question and to provide input on various potential study designs. This group of collaborators must assess the feasibility of study within the cooperative group system. Key issues such as the availability of cooperating investigators, the timeliness of the study, possible competing trials, regulatory requirements, and total cost are also considered.

The committee member who came up with the study idea is typically charged with functioning as national study chair of the protocol, though early-career investigators may be matched with a senior mentor. The national study chair will be responsible for developing the study as well as coordinating its conduct following activation. Co-chairs may be assembled from other members of the multidisciplinary team, as well as other committees. For example, quality of life co-chairs and imaging co-chairs may be responsible for the conduct of trial activities under these respective purviews. Co-chairs may also be assembled from other cooperative groups to facilitate intergroup collaboration.

Promising concepts are developed sequentially by the scientific committee over time based on discussions that are held both at the in-person group meetings and on teleconferences held between these major meetings. Because group meetings are typically held only twice a year, the design of a study may be a lengthy process.

17.4.2 Group and Task Force Review

Once the scientific committee has developed a broad idea and general plan for a study, it is formally drafted and submitted to the group's Study Concept Review Committee on an NCI/CTEP letter of intent (LOI) or concept submission form [31]. Fully developed concepts include scientific background, preliminary data, hypothesis, inclusion/exclusion criteria, and a plan for study implementation that includes its statistical design. Table 17.2 outlines the essential components of an LOI. Requests concerning study funding are also typically requested at this time. Ideally, submission to the Study Concept Review Committee takes place following approval of the

Table 17.2 Essential components of investigator's LOI criteria for NCI/CTEP evaluation

Rationale and background
Hypothesis
Eligibility criteria
Study design
Treatment plan
Correlative studies
Endpoints and statistical considerations
References

concept by the appropriate NCI Task Force and Steering Committee. Once approved, the concept is formally submitted by the group to the NCI's Cancer Therapy Evaluation Program (CTEP) for final approval.

17.4.3 Protocol Development

The study protocol is the formal plan for performing the trial. A well-designed study protocol results in efficient initiation and conduct of the trial and allows each participating investigator to better plan for staffing and implementation. Each participating investigator requires a protocol that is acceptable to his/her local institution and Institutional Review Board (IRB).

Following CTEP approval, the national study chair and other co-investigators finalize a draft of the protocol as well as construct the case report forms which will be used to collect data from each study participant. The study protocol and case report forms each then undergo a review process within the cooperative group. Once approved, these documents are submitted to the NCI. The protocol is, finally, forwarded to activation.

17.4.4 Study Activation

All studies conducted within the NCTN are available to all NCTN member sites. Following approval, the group's protocol office produces a memo that indicates a study is officially open for accrual. A formal announcement is made on the NCI website and the website of each of the participating cooperative groups. Once the trial has been activated nationally, it is available for local activation at any NCTN member site. Investigators at member sites decide which trials to activate locally on the basis of their institutions' research agenda and available resources. Investigators who wish to activate and register patients to a study must obtain approval from their IRB, although in certain situations and centers, approval from a central IRB may be sufficient.

Responsibility for the conduct of a study at the local level is assigned to a local primary investigator who must be credentialed and registered with the NCI. Funding for local conduct of the study is provided to each participating site through the cooperative group mechanism and is typically paid on a per-registration basis. These funds are used to pay for all study-related costs including nursing/oversight, regulatory, study drugs and procedures, etc. Certain studies may require additional credentialing and quality assurance. And, enrolling patients to a study may require enrollment to companion studies.

17.4.5 Study Monitoring

Study conduct requires significant effort from local investigators, regulatory committees at each participating institution, and the NCI. Data monitoring committees and multiple committees focused on quality control (e.g., pathology monitoring committee) may be utilized, both at the national and local level. These committees are independent of the investigators as they are designed to ensure both participant safety and study integrity. This requires periodic monitoring of data and performance [32]. As concerns for clinical trial integrity may occur [33–35], the independence of these groups is critical.

17.4.6 Why Become a Member of an NCTN Group?

Numerous reasons exist for surgeons to become involved as an engaged member of a network group (Table 17.3). First, there is inherent value in the ability to contribute meaningfully to surgical science. Involvement within the network groups also affords opportunities for networking with and receiving mentorship from leaders in the field. And, there are various avenues for leadership training and career development through scientific discovery and funding. Indeed, each of the cooperative groups offers multiple requests for application for funding throughout the year, some of which specifically target early-career investigators.

For new investigators, conducting cooperative trials may have advantages versus conducting institutional or industry trials. For example, cooperative group trials allow the study of larger, more diverse populations and offer the support of a significant regulatory and financial support infrastructure.

We feel, however, the best part of becoming involved is the fun! Participation in the system, though it may appear difficult to navigate, can significantly add to personal career satisfaction. Nonetheless, anyone responsible for conducting or participating in a cooperative trial through the NCTN should have a full understanding of the complexity of the undertaking. Many of the cooperative groups have early-career investigator committees or programs, which may hold new investigator courses to help new investigators navigate the cooperative group system [36].

Table 17.3 Personal advantages and disadvantages in being involved in an NCTN cooperative group

Advantages
Education
Networking and mentorship
Opportunity to work side-by-side with key opinion leaders
Discuss novel ideas and learn clinical trial development process
Leadership training and career development
Scientific contributions, publications
Opportunities for funding
Fun
Disadvantages
Process can be political
Time and effort
The system can be difficult to navigate

17.5 Conduct of Multicenter Trials Outside of the NCTN

The NCTN includes a well-funded research infrastructure that allows for the safe and efficient conduct of cooperative trials. However, conducting such trials without the support of the NCTN is certainly feasible. While many of the aforementioned steps required to develop and conduct a clinical trial are similar when performed outside of the NCTN, planning an MCT in this setting requires additional resources, an alternative infrastructure to deal with the regulatory burden involved, and a high degree of coordination among participating research sites. Planning for an MCT requires a similarly significant amount of preparation.

17.5.1 Regulatory Elements

Cooperative trials are extraordinarily complex. Financial teams are necessary to allocate and coordinate resources among participating sites. Standard operating procedures for research training, protocol implementation, patient enrollment, and treatment oversight across sites are necessary. Protocol and study support and management teams must exist to provide quality control mechanisms, to ensure high quality standards, and to confirm data integrity and compliance with regulatory requirements. Audit groups are necessary to evaluate the conduct of clinical trials and collaborate with all stakeholders to offer guidance and education. Centralized tissue banking and electronic health record (EHR) support may also be necessary to facilitate accrual, to promote patient safety, and to ensure quality biospecimen handling.

To deal with these and other complexities, major research institutions often have internal programs and teams designed to provide support to investigators engaged in

multicenter clinical trials. These programs provide support similar to those provided by the centralized infrastructure of the NCTN. If interested in conducting or participating in an MCT outside of the NCTN, it is important to identify any local support infrastructure and systems that may already be available within the institution or within the local healthcare network.

17.5.2 Feasibility and Costs

Cost is a critical concern for all clinical trials. Cost is a particular concern for MCTs given their typical complexity, size, and regulatory burden. Although funding for a cooperative trial may come from any type of sponsor, grants specifically designated to finance the planning of MCTs may exist from the NIH. Funds earmarked for the conduct of trials may also be awarded through the NIH or other sources.

Planning for a specific trial should place a primary emphasis on the assessment of feasibility. Pilot studies—small-scale tests of a planned larger study—may be useful in this regard. These studies may be used to provide baseline data and accurate sample size estimates, as well as estimates of the time and resources required to complete a larger, more statistically robust trial.

Within a study protocol, expenses can be minimized by limiting the number of scientific questions asked, by reducing the number of tests required as part of the protocol, by creating an efficient work and organizational structure, and by performing only necessary quality monitoring [37].

17.5.3 Consensus

Multiple institutions may be involved in cooperative trials, and these trials typically benefit from the expertise of a diverse group of investigators. However, achieving consensus among the group is critical, both to the design and conduct of research studies [38]. Reaching consensus may be difficult in many circumstances. Three popular and well-recognized methodologies may be used to promote consensus-building: focus group discussion [39], nominal group technique [40], and the Delphi method [41]. Each of these consensus-building methods has advantages and disadvantages. Ultimately, a combination of various methods for achieving group consensus can be employed.

17.5.4 Communication

Even if investigators think they have identified and dealt with all of the major scientific and logistical obstacles to the safe and efficient conduct of an MCT, unanticipated problems will invariably arise due to their size, complexity, and the number of

their associated investigators. This underscores the importance of effective communication. Study leaders must maintain open lines of communication with the participating institutions and investigators. Frequent in-person or virtual investigator meetings afford the ability to effectively communicate and address ongoing issues. Group meetings are also an ideal way to stimulate and maintain interest in a study.

17.6 Quality Control

Fundamental components of multidisciplinary cancer treatments include accurate preoperative disease staging, consistent and meticulous surgical technique, and precise pathologic analysis. The absence of standardization and quality control in any of these can adversely influence the accuracy of assessments of patients' eligibility for a clinical trial, the interpretation of the results of a trial, and ultimately, the outcomes of a trial. Quality assurance is therefore a critical area in the design of cooperative trial protocols.

In a recent analysis of patients enrolled in a national cooperative trial of postoperative therapy for pancreatic cancer, an overwhelming inconsistency in the surgical techniques used to resect study participants' tumors was identified [42]. Substandard surgical technique was hypothesized to contribute to the unexpectedly high rate of local recurrence observed following treatment with the study regimen. Variation in surgical quality has been demonstrated to influence rates of locoregional recurrence and long-term survival following surgery for rectal cancer [43–46]. Yet, in another report conducted by a cooperative group analyzing surgical and pathologic variables, surgeon variability and suboptimal surgical practices were identified in a large population of patients receiving protocol-based therapy [44]. In recognition of the potential influence of surgical technique on oncologic outcome and in an attempt to minimize variation, consensus guidelines for the perioperative management of patients with colorectal cancer have been established [47].

Although the College of American Pathologists (CAP) and the American Joint Committee on Cancer (AJCC) both have standard protocols in place to direct histopathologic analysis of surgical specimens resected in cancer operations, the degree to which they are observed varies [48–56]. Diversity among pathologists with regard to the methods used to analyze specimens may also help explain the disparities in rates of R1 resection and other important clinical metrics among clinical trial participants. A recent report has shown that direct interdisciplinary assessment instruments can be used to improve pathological quality assurance [57].

Finally, the importance of accurate perioperative staging must also be emphasized both as a means to predict prognosis and to accurately enroll patients in trials. Still, inconsistency has also been found in staging practice in a recent analysis [42].

Efforts to improve quality assurance within cooperative trials are ongoing. Emphasis on quality assurance was considered, for example, within the protocol of the recent "ALaCaRT" randomized clinical trial comparing laparoscopic-assisted versus open

resection in rectal cancer. There were strict criteria for surgeon eligibility and pathology assessment, and also audits of surgery, pathology, and other hospital data [58]. Eligibility criteria to identify surgeon qualification included evidence of laparoscopy expertise [58]. With regard to pathological assessment, all excision specimens were processed and analyzed according to protocol recommendations, and each pathologist was trained in assessing the mesorectal specimen [57, 59].

17.7 Final Words

In the end, cooperative trials play a major role in advancing knowledge in the fields of medicine and surgery. The NCI and its NCTN provide a robust infrastructure to help carry out these important studies for cancer, although MCTs certainly can be conducted outside of this mechanism with enough planning and coordination. In addition to the inherent value of the education and contribution to the scientific field, involvement in cooperative group research affords the opportunities for networking and mentorship, leadership training and career development, scientific discovery and funding, all of which can add to personal career satisfaction.

References

1. Juni P, Altman D, Egger M. Assessing the quality of randomised controlled trials. In: Egger M, Smith GD, Altman DG, editors. Systematic reviews in health care: meta-analysis in context. 2nd ed. London: BMJ Publishing Group; 2001.
2. Levin WC, et al. Cooperative clinical investigation. A modality of medical science. JAMA. 1974;227(11):1295–6.
3. Klimt CR. Principles of multi-center clinical studies. In: Boissel JP, Klimt CR, editors. Multicenter controlled trials. principles and problems. Paris: INSERM; 1979.
4. https://www.cancer.gov/research/key-initiatives/immunotherapy.
5. Coronary Drug Project Research Group. The Coronary Drug Project: methods and lessons of a multicenter clinical trial. Control Clin Trials. 1983;4(4):1–541.
6. Meinert CL. Clinical trials. Design, conduct and analysis. New York, NY: Oxford University Press; 1986.
7. Corless CL, et al. Pathologic and molecular features correlate with long-term outcome after adjuvant therapy of resected primary GI stromal tumor: the ACOSOG Z9001 trial. J Clin Oncol. 2014;32(15):1563–70.
8. Giuliano AE, et al. Effect of axillary dissection vs no axillary dissection on 10-year overall survival among women with invasive breast cancer and sentinel node metastasis: the ACOSOG Z0011 (Alliance) randomized clinical trial. JAMA. 2017;318(10):918–26.
9. Jang HJ, et al. The addition of bevacizumab in the first-line treatment for metastatic colorectal cancer: an updated meta-analysis of randomized trials. Oncotarget. 2017;8(42):73009–16.
10. Sandler A, et al. Paclitaxel-carboplatin alone or with bevacizumab for non-small-cell lung cancer. N Engl J Med. 2006;355(24):2542–50.
11. Early Breast Cancer Trialists' Collaborative Group, et al. Comparisons between different polychemotherapy regimens for early breast cancer: meta-analyses of long-term outcome among 100,000 women in 123 randomised trials. Lancet. 2012;379(9814):432–44.

12. http://iom.edu/Reports/2010/A-National-Cancer-Clinical-Trials-System-for-the-21st-Century-Reinvigorating-the-NCI-Cooperative.aspx.
13. Institute of Medicine. A national cancer clinical trials system for the 21st century: reinvigorating the NCI Cooperative Group Program. Washington, DC: The National Academies Press; 2010. https://doi.org/10.17226/12879.
14. https://www.ascopost.com/issues/march-15-2014/the-evolution-of-us-cooperative-group-trials-publicly-funded-cancer-research-at-a-crossroads/.
15. Patlak M, et al. Multi-center phase III clinical trials and NCI cooperative groups: workshop summary, vol. xii. Washington, DC: National Academies Press; 2009. p. 121.
16. https://dctd.cancer.gov/MajorInitiatives/NCI-sponsored_trials_in_precision_medicine.htm.
17. https://www.allianceforclinicaltrialsinoncology.org/main/.
18. https://ecog-acrin.org/.
19. https://www.nrgoncology.org/.
20. https://www.swog.org/.
21. https://childrensoncologygroup.org/.
22. https://www.ctg.queensu.ca/.
23. https://www.cancer.gov/research/nci-role/cancer-centers.
24. https://www.cancer.gov/research/areas/clinical-trials/nctn.
25. https://www.cancer.gov/research/areas/clinical-trials/ncorp.
26. https://www.cancer.gov/about-nci/organization/ccct/steering-committees/nctn.
27. https://www.cancer.gov/about-nci/organization/ccct.
28. https://deainfo.nci.nih.gov/advisory/ctac/ctac.htm.
29. American College of Surgeons Clinical Research Program, Alliance for Clinical Trials in Oncology, Nelson HD, Hunt KK. Operative standards for cancer surgery, volume i: breast, lung, pancreas, colon. Philadelphia, PA: Wolters Kluwer Health; 2015.
30. American College of Surgeons Clinical Research Program, Alliance for Clinical Trials in Oncology, Katz MHG. Operative standards for cancer surgery, volume ii: thyroid, gastric, rectum, esophagus, melanoma. Philadelphia, PA: Wolters Kluwer Health; 2018.
31. https://ctep.cancer.gov/protocolDevelopment/letter_of_intent.htm.
32. Friedman L, DeMets D. The data monitoring committee: how it operates and why. IRB. 1981;3:6–8.
33. Fleming TR, DeMets DL. Monitoring of clinical trials: issues and recommendations. Control Clin Trials. 1993;14(3):183–97.
34. Angell M, Kassirer JP. Setting the research straight in the breast-cancer trials. N Engl J Med. 1994;330:1448–50.
35. Cohen J. Clinical trial monitoring: hit or miss? Science. 1994;264(5165):1534–7.
36. Uy GL, Katz MHG, Boughey JC. Junior investigators: get engaged in the Alliance for Clinical Trials in Oncology. Bull Am Coll Surg. 2017;102(4):62–3.
37. Eisenstein EL, et al. Sensible approaches for reducing clinical trial costs. Clin Trials. 2008;5(1):75–84.
38. Chung KC, Song JW, W.S. Group. A guide to organizing a multicenter clinical trial. Plast Reconstr Surg. 2010;126(2):515–23.
39. Krueger RA, Casey MA. Focus groups: a practical guide for applied research. Thousand Oaks, CA: Sage Publications; 2006.
40. Van de Ven AH, Delbecq AL. The nominal group as a research instrument for exploratory health studies. Am J Public Health. 1972;62:337–42.
41. Dalkey NC. The Delphi method: an experimental study of group opinion. Santa Monica, CA: RAND Corp; 1969.
42. Katz MH, et al. Standardization of surgical and pathologic variables is needed in multicenter trials of adjuvant therapy for pancreatic cancer: results from the ACOSOG Z5031 trial. Ann Surg Oncol. 2011;18(2):337–44.
43. Phillips RK, et al. Local recurrence following 'curative' surgery for large bowel cancer: I. The overall picture. Br J Surg. 1984;71(1):12–6.

44. Stocchi L, et al. Impact of surgical and pathologic variables in rectal cancer: a United States community and cooperative group report. J Clin Oncol. 2001;19(18):3895–902.
45. Holm T, et al. Influence of hospital- and surgeon-related factors on outcome after treatment of rectal cancer with or without preoperative radiotherapy. Br J Surg. 1997;84(5):657–63.
46. Kockerling F, et al. Influence of surgery on metachronous distant metastases and survival in rectal cancer. J Clin Oncol. 1998;16(1):324–9.
47. Nelson H, et al. Guidelines 2000 for colon and rectal cancer surgery. J Natl Cancer Inst. 2001;93(8):583–96.
48. American Joint Committee on Cancer. Exocrine pancreas. In: Greene FL, Page DL, Fleming ID, et al., editors. AJCC cancer staging manual. 6th ed. Chicago, IL: Springer; 2002. p. 157–64.
49. College of American Pathologists. Cancer protocols [16 August 2010]; Pancreas (exocrine); 2009. http://www.cap.org/apps/docs/committees/cancer/cancer_protocols/2009/PancreasExo_09protocol.pdf.
50. American Joint Committee on Cancer. Exocrine pancreas. In: Edge SB, Byrd DR, Compton CC, et al., editors. AJCC cancer staging manual. 7th ed. Chicago, IL: Springer; 2009. p. 241–9.
51. Menon KV, et al. Impact of margin status on survival following pancreatoduodenectomy for cancer: the Leeds Pathology Protocol (LEEPP). HPB (Oxford). 2009;11(1):18–24.
52. Staley CA, et al. The need for standardized pathologic staging of pancreaticoduodenectomy specimens. Pancreas. 1996;12(4):373–80.
53. Chatelain D, Flejou JF. Pancreatectomy for adenocarcinoma: prognostic factors, recommendations for pathological reports. Ann Pathol. 2002;22(5):422–31.
54. Luttges J, Zamboni G, Kloppel G. Recommendation for the examination of pancreaticoduodenectomy specimens removed from patients with carcinoma of the exocrine pancreas. A proposal for a standardized pathological staging of pancreaticoduodenectomy specimens including a checklist. Dig Surg. 1999;16(4):291–6.
55. Khalifa MA, Maksymov V, Rowsell C. Retroperitoneal margin of the pancreaticoduodenectomy specimen: anatomic mapping for the surgical pathologist. Virchows Arch. 2009;454(2):125–31.
56. Verbeke CS, et al. Redefining the R1 resection in pancreatic cancer. Br J Surg. 2006;93(10):1232–7.
57. Nagtegaal ID, et al. Macroscopic evaluation of rectal cancer resection specimen: clinical significance of the pathologist in quality control. J Clin Oncol. 2002;20(7):1729–34.
58. Stevenson AR, et al. Effect of laparoscopic-assisted resection vs open resection on pathological outcomes in rectal cancer: the ALaCaRT randomized clinical trial. JAMA. 2015;314(13):1356–63.
59. https://www.rcpa.edu.au/Library/Practising-Pathology/Structured-Pathology-Reporting-of-Cancer/Cancer-Protocols.

Chapter 18
International Trials: Surgical Research Networks

Marc A. Gladman

18.1 Introduction

In 1996, Richard Horton, the editor of the *Lancet*, published a damning commentary "Surgical Research or Comic Opera: questions, but few answers" strongly criticizing the poor quality of surgical evidence and the lack of clinical trials [1]. Having reported that almost half of all publications were in the form of case series, he called into question the value of a large proportion, and indeed the future, of surgical research and insisted that surgeons needed to find ways to conduct clinical trials to retain their academic reputation [1]. In the two decades that have followed, surgeons across the globe have responded, and journals are now populated with high-quality surgical trials. Internationally, two important initiatives have facilitated the conduct of such trials. First, instead of conducting surgical trials across single or multiple regional centers, networks spanning entire nations/continents have been established [2]; remarkably, the trials across these networks are run by surgical trainees and medical students. Second, surgical trials have gone "global" and international collaborative surgical trials networks have been established [3]; importantly, centers from low- and middle-income countries (LMICs) are included in these networks. Today, the future of surgical research appears assured and the academic reputation of surgery is preserved. This chapter will describe these national, continental, and international networks, highlighting the key organizational factors crucial in their establishment to facilitate the conduct of high-quality clinical trials in surgery.

M. A. Gladman (✉)
Adelaide Medical School, Faculty of Health and Medical Sciences, The University of Adelaide, Adelaide, SA, Australia
e-mail: marc.gladman@adelaide.edu.au

T. M. Pawlik, J. A. Sosa (eds.), *Clinical Trials*, Success in Academic Surgery,
https://doi.org/10.1007/978-3-030-35488-6_18

18.2 What Are Surgical Trials Networks and What Are Their Benefits?

Clinical trials, involving the allocation of human participants to one or more health-related interventions to evaluate the effects on health outcomes [4], are crucial in assessing the safety, efficacy, effectiveness, and acceptability of surgical interventions. The obvious challenges faced when planning/conducting clinical trials in surgery are well documented and include difficulties with respect to blinding, surgical disease, and intervention complexity, inconsistent expertise of care providers, and differences in surgical volume and hospital resources and health system processes between different centers [5, 6]. Accordingly, less than 1% of patients have historically been enrolled in surgical clinical trials [3]. However, this situation is changing and, rather than working in isolation in single-centers, individual surgeons are linking together with colleagues to form research collaboratives to deliver multi-center trials.

Performing small studies in individual units is unlikely to achieve high-impact publication and change surgical practice due to design limitations (small numbers, short follow-up, etc.), even if the clinical questions being addressed are valid. By contrast, working together and conducting trials in an identical manner across multiple units simultaneously allows the inclusion of a larger number of participants, across a wider range of population groups with the ability to compare results among centers in different geographic locations [7]. This in turn leads to the acquisition of more robust data that is more likely to be published in high-impact journals and influence patient care. The establishment of a network of collaborating units into a "surgical trials network" with inclusion of centers spanning entire regions, states, countries, continents, and even the globe enhances the subsequent generalization of the trial findings [8]. Such networks now exist in the United Kingdom, Australia and New Zealand, and pan-Europe (see below). Important benefits are achieved by "formalizing" these networks and have helped tackle some of the complexities of conducting trials across multiple centers by providing efficient central coordination and control of trials that facilitate standardized protocol implementation, data management, and quality assurance.

Whilst the creation of national surgical trial networks has been an important development in surgical research, the burden of surgical disease is global, with an estimated 235 million major surgical procedures performed annually [9]. Historically, however, few clinical trials have involved international collaborations [3]. In 2015, the Bangkok Global Surgery Declaration called upon the world to "Strengthen Emergency and Essential Surgical Care and Anesthesia as a part of Universal Health Coverage" and committed to "support activities that promote global collaboration among all countries and regions to work towards implementation solutions for ensuring universal access to safe, affordable surgical and anesthesia care when needed" [10]. Increasingly, LMICs report similar surgically treatable diseases (i.e., cancer, cardiovascular disease etc.) as developed countries but the few international surgical trials that have been conducted have failed to include LMICs,

despite the obvious mutual interest of research collaboration [3]. The potential for major public health gains through international collaboration in surgical research is high and extending trials to LMICs allows a greater number of trials to be conducted and increases the generalizability of the results achieved. GlobalSurg and the National Institute of Health Research (NIHR) Unit on Global Surgery in the United Kingdom (see below) are important recent initiatives that have facilitated the creation of international collaborative networks with the deliberate strategy of including LMICs.

18.3 Nationwide and Pan-Continental Surgical Trials Networks

18.3.1 The U.K. Surgical Trials Initiative

In 2011, The Royal College of Surgeons in England, together with the National Institute of Health Research, Cancer Research U.K. and charitable partners, established a network of surgical trial units across the United Kingdom to deliver multicenter clinical trials within surgical specialties. Over 65 trials have been registered to date, with active recruitment into over 40 trials, but the ultimate ambitious aim of this initiative is to allow any surgical patient in the United Kingdom who wishes to join a trial will be able to do so, and to ensure that every surgical trainee, by the time they become a consultant (attending), will know that conducting trials is an essential part of consultant (attending) activity and not be an optional extra [2]. Central to this initiative was the establishment of a number of designated surgical trials centers across the country with expertise in study design, data collection, and analysis. Surgical specialty leads were appointed to work closely with these centers to develop clinical networks and ensure delivery of the trials. Extraordinarily, the trials are led by surgical trainees (residents), and regional networks in general surgery have been developed to adopt a novel collaborative approach to research in the United Kingdom (Fig. 18.1) [11].

18.3.1.1 Can Surgical Trainees Really Run Complex Clinical Trials in Surgery?

As an apparent world-first, the central strategy (and key ingredient to success) of the surgical trials initiative in the United Kingdom has been the trainee-led model of surgical research networks. All collaboratives are aligned with existing organized structures, including professional specialty associations (Fig. 18.1), which in the United Kingdom includes the Association of Surgeons in Training, the Royal College of Surgeons of England, and the Association of Surgeons of Great Britain and Ireland who provide academic, structural, and logistical support to surgical

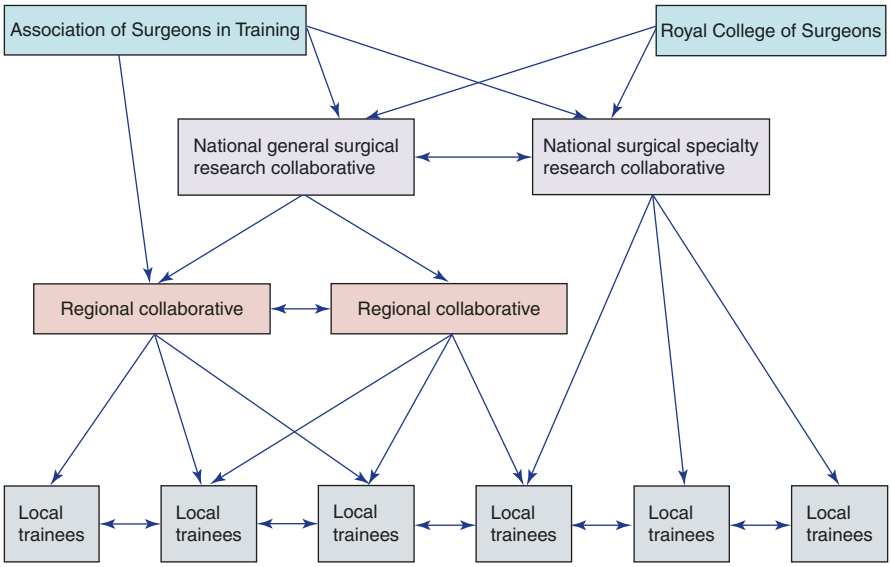


Fig. 18.1 Organization of the U.K, Trainee-led Research Collaborative Network. The Royal College of Surgeons and the Association of Surgeons in Training oversee a national general surgical and surgical specialty research collaborative. The trials are led by trainees in networks supported by Surgical Trials Centers and Surgical Specialty Leads. Reproduced from: Bhangu N, et al. Surgical research collaboratives in the UK. The Lancet 2013;382(9898):1091–1092. Copyright © 2013 Elsevier Ltd

trainee collaborative development. However, is this support enough to allow trainees to successfully complete complex trials? The first regionally developed general surgical research collaborative was the West Midlands Research Collaborative (WMRC) who performed the ROSSINI (Reduction Of Surgical Site Infection using a Novel Intervention) randomized controlled trial that recruited 760 patients across 21 centers to use either a wound-edge protection device or standard practice during wound closure [12]. The rapid recruitment, large number of patients (providing adequate power), impeccable organization and coordination, and minimal loss to follow-up demonstrated the ability of trainees to plan, conduct, and publish high-quality multicenter research in high-impact peer-reviewed journals [12].

It is evident that these collaboratives are focused on designing trials to actually improve surgical care rather than just accumulating research outputs. The recently registered follow-up ROSSINI-2 study will evaluate the use of three "in-theatre" interventions to reduce SSI rates in patients undergoing surgery involving an abdominal incision. Importantly, and in an attempt to deliver the highest quality trials possible, as the experience and success of the trainees within the research collaborative have increased, so has the complexity of the trial design. For example, the trainee-led ROSSINI-2 that has just launched is a phase III, multi-arm, multi-stage (MAMS) pragmatic, blinded (patient and outcome assessor) multicenter, randomized controlled trial with an internal pilot with a non-factorial superiority design with

allocation of various combinations of the three interventions to be used during the same operation, via seven possible treatment arms plus one control arm.

Following the success of the WMRC, ten diverse surgical research networks were subsequently established allowing for almost complete coverage of the United Kingdom. This National Surgical Research Collaborative went on to deliver the Multicenter Appendicectomy Audit, including 3326 consecutive patients undergoing appendicectomy from 95 centers during a 2-month period to investigate variation in provision and outcome of emergency appendicectomy [13]. Clearly, this demonstrates that surgical trainees can lead trials not only at regional but also at national level. However, "going national" requires additional structural considerations to ensure successful implementation across large geographic areas. To facilitate the planning and communication, networks were organized by surgical specialty to plan randomized trials focused on important specialty-specific research questions. For example, in gastrointestinal surgery, the DREAMS trial investigated dexamethasone versus standard treatment for postoperative nausea and vomiting in gastrointestinal surgery [14], and the ROCSS (Reinforcement of Closure of Stoma Site) study [15] involved a successful collaboration between the trainee network and industry to complete a large randomized trial for a complex surgical intervention. The model has also worked for more specialized surgical disciplines (e.g., pediatric surgery, neurosurgery, plastic surgery, and cardiothoracic surgery) that are centralized in tertiary units or large regional hospitals, and so require a national rather than regional network (Fig. 18.1). In 2012, the British Neurosurgical Trainee Research Collaborative was founded [16], the organization of which is shown in Fig. 18.2. In 2013, it launched its first national study: a prospective cohort study of patients with chronic subdural hematoma, which collected data from 1205 patients from 26 units across the United Kingdom [17] and has since launched more than 10 projects, including randomized trials and further prospective cohort studies [16].

So, do these collaboratives actually cover the entire nation? In general surgery alone, 238 of 241 U.K. hospitals (99%) providing general surgery services have participated in one or more trainee-led collaborative studies over the past decade [8]. Trainee groups have successfully delivered 15 studies: 12 observational studies and three randomized controlled trials, coordinated by five regional and two national trainee networks [8]. Examples of the general surgical trials conducted are presented in Table 18.1. While the possibility of surgical trainees leading trials a decade ago was almost inconceivable, the incredible success and achievements of the U.K. network has overwhelmingly proven that anything is possible. Indeed, trainees are ideally placed to deliver this model as they are motivated and are keen to engage in formalized research and audit activities. Importantly, they are in regular contact with each other, increasingly via social media, and work at the coalface in direct contact with patients facilitating their recruitment and enabling the engagement of all surgical units. In the United Kingdom, surgical trainees usually follow a rotational pattern through several hospitals in a region of the country. Even though this is not always the case in other countries, young trainee connectedness and readiness can facilitate the formation of research networks anywhere in the globe.

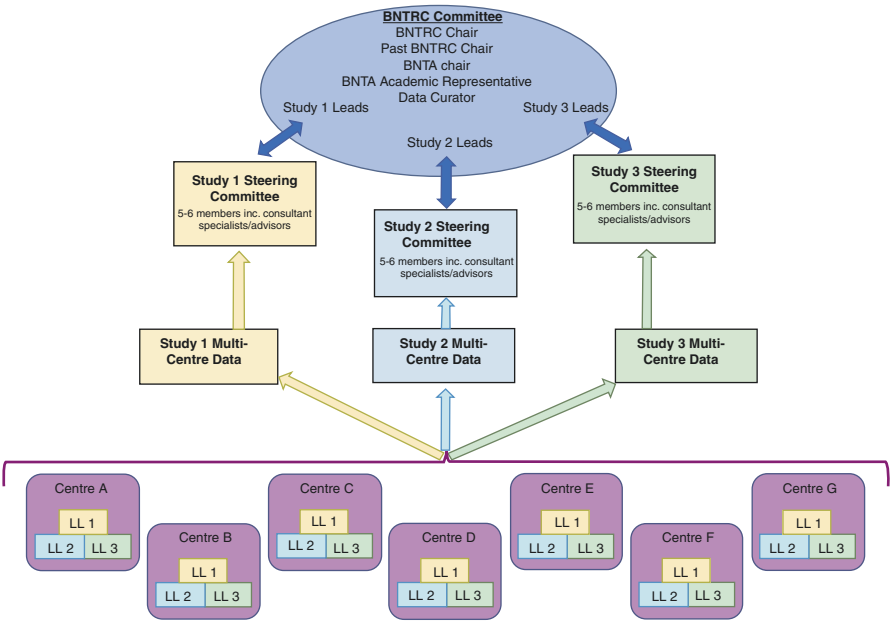


Fig. 18.2 Organization of The British Neurosurgical Trainee Research Collaborative (BNTRC). The BNTRC committee is composed of the individual study leads and representatives from the British Neurosurgical Trainee Association (BNTA). Each study has a study steering committee and local leads (LL) at each neurosurgical center. Reproduced from: Chari A, et al. The British Neurosurgical Trainee Research Collaborative: Five years on. Acta Neurochir (Wien). 2018;160(1):23–28. Copyright © 2017, The Author(s)

18.3.2 The Clinical Trials Network Australia and New Zealand

Having witnessed the remarkable success in the United Kingdom, The Royal Australasian College of Surgeons (RACS), through its Section of Academic Surgery, was keen to try to emulate the model to set up a clinical trials network in Australia and New Zealand. This was all the more challenging given the enormous land mass and geographic divide between two neighboring countries and the fact that surgical training is overseen by totally separate bodies within each of these two countries. In 2017, the Clinical Trials Network Australia New Zealand (CTANZ) was established following close communication with the Directors of the UK Surgical Trials Initiative. CTANZ seeks to inspire current and future trainees to make a difference by incorporating research into their daily practice. As in the U.K. model, surgical trainees working in specialty networks are the principal investigators recruiting patients into multi-center, prospective, clinical trials. The emphasis is on empowering surgical trainees to design, conduct, analyze, and publish clinical trials during their training. Again, the trainee networks are supported by appointed Surgical Specialty Leads who mentor and provide arm's length guidance to the networks. CTANZ has worked closely with the RACS Trainee Association (RACSTA) to

Table 18.1 Output of U.K. trainee research collaboratives in general surgery

Study	Coordinating group	Data collection	UK centers included	Status
National Appendicitis Audit	WMRC	2012	76	Published [13]
Survey of Hernia Antibiotic Prophylaxis usE (SHAPE)	LSRG	2012	34	Published [33]
National Sepsis Audit	SPARCS	2013	100	Published [34]
STARSurg-1	STARSurg	2013	108	Published [18]
Complicated Acute Diverticulitis Study (CADS)	YSRC	2014	105	Analysis in progress
Determining Surgical Complications in the Overweight (DiSCOVER)	STARSurg	2014	151	Published [19]
CholeS	WMRC	2014	150	Published [35]
Packing of Perianal Abscess Cavities (PPAC)	NWRC	2014	12	Published [36]
GlobalSurg 1	GlobalSurg	2016	98	Published [30]
Outcomes After Kidney injury in Surgery (OAKS)	STARSurg	2015	160	Analysis in progress
EuroSurg-1	EuroSurg	2016	14	Analysis in progress
GlobalSurg 2	GlobalSurg	2016	44	Analysis in progress
Reduction of Surgical Site Infection using a Novel Intervention (ROSSINI)	WMRC/BCTU	2010–2011	21	Published [12]
Dexamethasone Reduces Emesis After Major gastrointestinal Surgery (DREAMS)	WMRC/BCTU	2011–2014	45	Published [14]
Reinforcement of Closure of Stoma Site (ROCSS)	WMRC/BCTU	2012–2017	32	Published [15]

LSRG London Surgical Research Group, *NWRC* North West Research Collaborative, *SPARCS* Severn and Peninsula Audit and Research Collaborative, *STARSurg* Student Audit and Research in Surgery Collaborative, West Midlands Research Collaborative, *YSRC* Yorkshire Surgical Research Collaborative
BCTU Birmingham Clinical Trials Unit, *CHaRT* Centre for Healthcare Randomised Trials, CTRU-S Clinical Trials Research Unit, University of Sheffield, *QMUL* Queen Mary University of London

establish seven trainee-led networks (see Table 18.2). These networks have contributed and are already contributing to numerous trials including: PREVENTT (Preoperative Intravenous iron for anemia before major abdominal surgery); ITACS (Intravenous iron to treat anemia before cardiac surgery) ; ECST-2 (European Carotid Surgery Trial 2); POISE-3 (PeriOperative ISchemic Evaluation-3); GIVE (Groin wound Infection after Vascular Exposure); SUNRRISE (Single Use Negative pRessure dressing for Reduction In Surgical site infection following Emergency laparotomy); and IMAGINE (Ileus Management International).

Table 18.2 CTANZ trainee-led research networks

Name of network	Span of network
General surgery	
VERITAS Victorian Collaborative for Education, Research, Innovation, Training and Audit by Surgical Trainees	Victoria, Tasmania, Northern Territory
QUEST Queensland Surgical Trainee Research Collaboration	Australia and New Zealand
STARC South Australian Trainees Audit & Research Collaborative	South Australia
STORCC Surgical Trainee Organisation for Research, Central Coast	Central Coast, New South Wales
Orthopedic surgery	
New Zealand Orthopedic Surgery Network	New Zealand
Pediatric surgery	
ANZSCRAFT	Australia and New Zealand
Australia and New Zealand Surgery in Children Registrar's Association For Trials	
Vascular surgery	
Perth Clinical Trials Unit	Australia, UK, Europe, Asia

18.3.3 STARSurg: Whatever Surgical Trainees Can Do, so Can Medical Students!

Inspired by the efforts of their recently graduated colleagues entering surgical training, medical students in the United Kingdom with an interest in pursuing a career in surgery refused to be outdone. The Student Audit and Research in Surgery (STARSurg) collaborative is a national, student-led audit and research network with representation from medical schools across the United Kingdom and Ireland. It empowers students to participate in high-quality academic projects, forming links with supervising trainee and consultant (attending) surgeons. But if surgical trainees conducting trials seemed unlikely, surely the prospect of medical students contributing to, running and publishing, surgical trials is an impossibility. Apparently not. STARSurg's first national project saw 258 student collaborators representing 31 U.K. medical schools work together to collect outcomes data over a two-week period on a prospective cohort of 1500 patients across 109 U.K. hospitals that went on to be published in Europe's leading surgical journal [18]. To prove that this wasn't "beginner's luck" they subsequently went on to publish prospective, observational, multicenter studies investigating the impact of body mass index on postoperative complications following gastrointestinal surgery [19] and acute kidney injury following major elective non-cardiac surgery [20, 21]. RECON (REspiratory ComplicatiOns after abdomiNal Surgery) is the latest STARSurg international audit and will investigate the inci-

dence of postoperative pulmonary complications (PPCs) following major abdominal and incisional hernia surgery and evaluate adherence to perioperative measure to reduce their risk.

18.3.4 EuroSurg: The European Student Research Collaborative

The EuroSurg collaborative is a pan-European medical student and trainee-led surgical research network that runs high-quality, multi-center, international studies. It adopts a similar organizational structure to STARSurg (see Fig. 18.3). Data collection is performed by "mini teams" of two or three medical students and at least one junior doctor (resident), supervised by a senior consultant (attending) throughout the data collection period. Local leads are in charge of organizing students and doctors into these mini teams and are responsible for the smooth running of the project in individual medical schools. Finally, regional leads ensure the overall running of the project in respective countries. They are responsible for coordinating with local

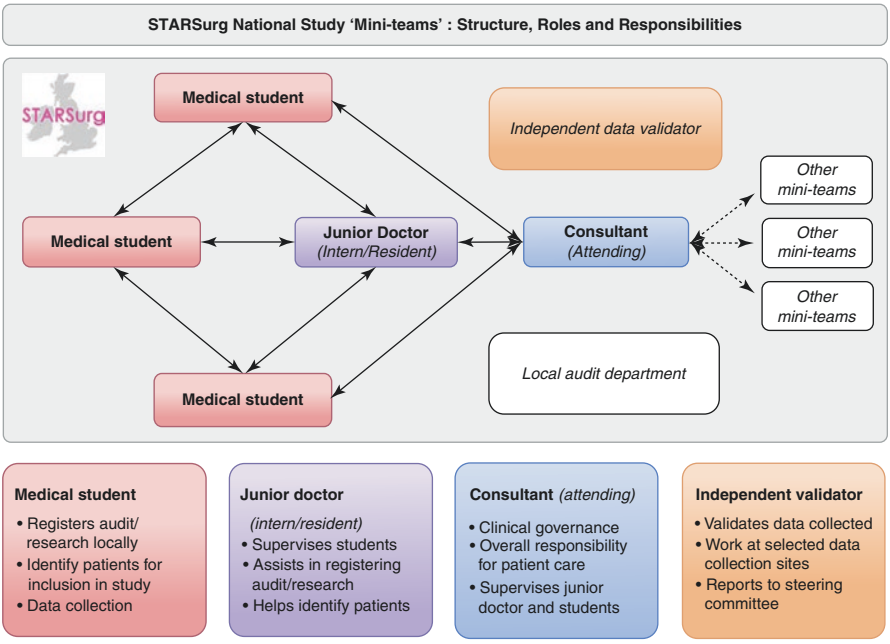


Fig. 18.3 Organization of the medical student-led research collaborative – STARSurg / EuroSurg: Reproduced from: STARSurg Collaborative. Outcomes After Kidney injury in Surgery (OAKS): protocol for a multicentre, observational cohort study of acute kidney injury following major gastrointestinal and liver surgery. BMJ Open. 2016 Jan 14;6(1):e009812. Copyright © 2016 with permission from BMJ Publishing Group Ltd

leads to ensure the project is running well in each medical school. They also update the steering committee on the progress of the project. In 2018, EuroSurg conducted IMAGINE (Ileus Management International), a multicenter prospective cohort study across two continents: Europe and Australasia [22]. This study aimed to assess the role of non-steroidal anti-inflammatory drugs (NSAIDs) for reducing ileus after surgery, recruiting over 4000 patients and has recently been published [23].

18.4 Global Surgical Trials Networks

International collaboration in surgical research, facilitating large multinational trials that cross cultures and levels of socioeconomic development, should have faster results and wider applicability than single-country trials. However, conducting trials across multi-centers in different countries is complex and requires painstaking co-ordination, quality control, and data management. Furthermore, the manner in which the protocol is implemented should be clear and similar at all centers. Despite these apparent disincentives, there are examples of high-quality surgical trials being successfully conducted by international collaborators, one of which was the STICH randomized trial that compared early surgery with initial conservative treatment for patients with intracerebral hemorrhage, recruiting over 1000 patients from 83 centers in 27 countries [24]. Currently, there are numerous international surgical trials being conducted, some examples of which are presented in Table 18.3.

Given that the burden of surgical disease is global, and that LMICs report similar surgically treatable diseases as developed countries, there is obvious mutual interest for research collaboration and the potential for major public health gains is high [3]. In the 1980s, HIV surged onto the public health agenda and gained international coverage. During the ensuing two decades, the number of global clinical drug trials expanded with developing countries increasingly involved due to lower costs, improved access to previously untreated patients, and improvements in healthcare infrastructure in these regions. Accordingly, considerable progress has been made tackling many non-surgical global diseases. Less encouragingly, it has recently

Table 18.3 International collaborative surgical trials

ROLARR: Robotic versus laparoscopic surgery for rectal cancer surgery
GLiSten: Next-Generation intraoperative Lymph node staging for Stratified colon cancer surgery
LAVA: Liver resection surgery versus thermal ablation for colorectal LiVer MetAstases
IntAct: Intraoperative Fluorescence Angiography to Prevent Anastomotic Leak in Rectal Cancer Surgery
COMICS: Conventional versus Minimally Invasive extra-corporeal circulation in patients undergoing Cardiac Surgery: a randomized controlled trial
VERDICT: Preoperative Volume Replacement vs. usual care in Diabetic patients having CABG surgery: a randomized controlled trial
STAR-TREC: *T*rans*A*nal microsurgery following (chemo) *R*adiotherapy versus *T*otal mesorectal excision for early *RE*ctal *C*ancer

been demonstrated that more people (4.2 million) die within 30 days of surgery each year than from all causes related to HIV, malaria, and tuberculosis combined (2.97 million deaths) [25]. Indeed, this number of postoperative deaths accounts for 7.7% of all deaths globally, making it the third greatest contributor to deaths, after ischemic heart disease and stroke [25]. Furthermore, five billion people do not have access to safe, timely, and affordable surgical care. In LMICs, the problem is even more acute, with 9 out of 10 people lacking access to even the most basic surgical services.

Unfortunately, the worldwide expansion of surgical trials has only occurred this decade [3]. While considered ambitious, or perhaps even impossible, the successful completion of trials emanating from international research collaboration across countries of different socioeconomic development confirms that such an aspiration can be achieved by surgeons. Examples include the WHO checklist development in the Safe Surgery Saves Lives project [26], the CRASH-2 (Clinical Randomization of an Antifibrinolytic in Significant Hemorrhage) trial on tranexamic acid in trauma [27, 28], and development of globally agreed metrics of outcome surveillance [29].

18.4.1 The Global Surgical Outcomes Collaborative: GlobalSurg

GlobalSurg was founded in 2013 and is a collaborative of surgeons and research methodologists who aim to conduct pragmatic, patient-facing research focused on LMICs. GlobalSurg has successfully completed two observational studies in abdominal surgery, involving 25,000 patients from over 100 countries. GlobalSurg-1 aimed to identify variation in outcome of emergency intra-abdominal surgery across international settings and demonstrated that 24-h and 30-day mortality were higher in LMICs compared to high-income countries [30], and GlobalSurg-2 aimed to determine worldwide surgical site infection (SSI) rates following gastrointestinal surgery and demonstrated that they were twice as common in LMICs [31]. Currently, GlobalSurg-3 is collecting data and aims to determine variation in quality of cancer surgery worldwide by measuring 30-day mortality and complication rates in patients undergoing surgery for breast, gastric, and colon cancers.

18.4.2 The UK NIHR Global Health Research Unit on Global Surgery

The NIHR Global Health Research Unit on Global Surgery was established in 2017 and is led by the University of Birmingham in the United Kingdom, in partnership with the Universities of Edinburgh and Warwick, along with partners from the GlobalSurg Collaborative in a number of LMICs. The main objective of the unit is

to build sustainable clinical research capacity in LMICs by establishing independent research hubs to conduct clinical trials for surgical patients. Each hub runs clinical trials and cohort studies and supports research training and education. In addition, the hub center works with other "spoke" hospitals within their country, supporting them to conduct trials.

The unit is currently coordinating three clinical trials. The Falcon trial has taken the baseline data provided by GlobalSurg-2 and is a pragmatic multi-center factorial randomized controlled trial testing measures to reduce surgical site infection in LMICs. Following GlobalSurg-3, the Crane trial will be the first high-quality global cluster randomized trial of a nutritional supplement given to malnourished patients prior to cancer surgery in LMICs. The European Society of Coloproctology safe-anastomosis program (EAGLE) is an international, cluster randomized-sequence study of a safe-anastomosis quality improvement intervention to reduce anastomotic leak following right colectomy and ileocaecal resection and will collaborate with researchers in high-, middle-, and low-income countries all around the world, involving hospitals across Europe, South East Asia, South America, the United States, Saudi Arabia, North Africa, and Russia.

18.5 Establishing Surgical Research Collaborative Networks: Practical Considerations

Getting started is always the biggest challenge. Trainee-led research collaboratives, as described in this chapter, are complex groups involving multiple units, each of which are centered in individual hospitals in geographical regions. Each of these has a degree of academic heterogeneity in terms of resources that may include trialists, epidemiologists, health economists, and clinicians, working jointly from discipline-specific bases. Naturally, the most important determinant of success is teamwork and effective communication between all stakeholders involved. The key groups of individuals to identify and involve at inception are:

1. Enthusiastic, committed local surgical trainees (residents) at each site to lead the trials; clinical (for ideas generation) and local (logistics of hospitals, etc.) knowledge are essential.
2. Senior surgeons (attendings) at each of the sites to mentor, support, and advocate for the junior trainees locally.
3. Trial experts who can advise regarding study design, methodology, statistics, and analysis.

Usually, interest is not a problem, but it is important to "work with the willingness" from the outset and identify "champions" who will advocate and support the development of the collaboration. Next, an organizational structure needs to be established to allow "control and command" of the projects/trials. The networks presented in this chapter have all adopted a very similar structure that has proved successful and achieved reliable outputs. Typically, this involves three tiers:

1. *An expert steering committee* that is involved in protocol design and initiation of communication centrally to teams and trainees.
2. *A multidisciplinary team involving trial experts/senior surgeons* (usually specialty led) with experience conducting trials. This team provides academic support and organizes/communicates with the local teams.
3. *Local (hospital-level) teams of several trainees (residents) and/or medical students* paired with a consultant (attending) surgeon for advocacy and mentorship. The local teams are responsible for patient recruitment and data collection/entry.

The organizational "hierarchy" presented above should be meant for the purposes of "direction/steering" only, as it is important to note that trainees are the critical element of success and for the trials to be run with a "bottom-up" trainee (resident)-driven rather than "top-down" consultant (attending) dictated approach. The trainees need to be empowered and allowed to function autonomously. However, they need access to senior clinicians/academics with experience conducting trials for support, mentorship, and academic leadership when appropriate, despite being trainee-led.

The last component to consider is the operationalization of the trial network. For early promotion of the concept and engagement of trainees, an "old-fashioned" face-to-face forum, preferably a break-out from other existing scheduled activities that brings trainees together (e.g., training meetings/scientific congress), is a great way to achieve this when establishing the new network. Communication and information sharing are fundamental. Continued communication between trainees and senior surgeons/mentors and, perhaps more importantly, between trainees, is crucial. Whilst this can be promoted and enhanced using modern technology and social media, regular meetings/forums can serve as a good opportunity for sharing of ideas/problems/solutions. Scheduling is always a challenge but predictable, accessible meetings that create minimal inconvenience for all concerned tend to be most fruitful. Early on, trainees will need more support/access to the expertise of senior surgeons/specialty leads. Starting with simple trials to cultivate interest and build confidence before working up to more ambitious projects will ensure the creation of a robust network and will translate rapidly to up-skilling of trainees and research outputs. Dissemination of information between the "hierarchical levels" and clear definition of roles and goals for all concerned will lead to efficient operations.

Important principles for the ongoing management of research collaboratives so that they deliver high-quality research projects have been shared by the West Midlands Research Collaborative (WMRC); the first trainee-led collaborative that have produced sustained outputs for over a decade and include: (1) engagement of committed trainees who are determined and can corral colleagues into contributing; (2) ensuring shared benefits by crediting all contributors in final publications; (3) obtaining endorsement and encouragement from national surgical/research organizations/institutions; (4) enlisting inspirational senior mentors; (5) retaining active trainee-level leadership so that committee members build experience; (6) developing the network locally and expand it as it becomes more established; (7) identifying supportive academics to gain partnership for funding applications; and (8) delivering efficient administration (central trials office and at each local site, etc.) [32].

Important considerations for sustainability and longevity are authorship, research outputs, and funding. All those actively involved in collaborative trials and studies are credited in final papers in the collaborative models presented in this chapter in line with indexing and journal policies for large studies. Authorship policies should be agreed upon at the start of all projects and strictly adhered to. Research outputs take time to emerge but success breeds success and help build a track-record for future funding applications. Obtaining funding for projects can be challenging, but linking trainees with academics who have secured competitive funding for guidance and mentorship gets them involved in the process.

18.6 Conclusion

During the last two decades, the quality of surgical research has improved dramatically. Surgeons have become engaged in the conduct of clinical trials across multiple centers, which has been facilitated by the creation of trainee-led networks that span entire nations, continents, and now even the globe—with the inclusion of LMICs. For the population, this should lead to improved care for surgical patients, the global delivery of advances in surgical care, and dissemination of best practice. For the profession, a culture of trials will be established in surgical practice as these trainees become consultants (attendings). Surgeons will be trained to enter patients in clinical trials, which will hopefully become the "norm" and an "essential" rather than a "desirable" attribute for a consultant (attending) surgeon. Contrary to previously articulated concerns, the future of surgical research appears assured and the academic reputation of surgery preserved.

References

1. Horton R. Surgical research or comic opera: questions, but few answers. Lancet. 1996;347(9007):984–5.
2. McCall B. UK implements national programme for surgical trials. Lancet. 2013;382(9898):1083–4.
3. Soreide K, Alderson D, Bergenfelz A, Beynon J, Connor S, Deckelbaum DL, et al. Strategies to improve clinical research in surgery through international collaboration. Lancet. 2013;382(9898):1140–51.
4. Laine C, Horton R, DeAngelis CD, Drazen JM, Frizelle FA, Godlee F, et al. Clinical trial registration—looking back and moving ahead. N Engl J Med. 2007;356(26):2734–6.
5. Weil RJ. The future of surgical research. PLoS Med. 2004;1(1):e13.
6. Khajuria A, Agha RA. Surgical clinical trials—need for quantity and quality. Lancet. 2013;382(9908):1876.
7. Appel LJ. A primer on the design, conduct, and interpretation of clinical trials. Clin J Am Soc Nephrol. 2006;1(6):1360–7.
8. Nepogodiev D, Chapman SJ, Kolias AG, Fitzgerald JE, Lee M, Blencowe NS, et al. The effect of trainee research collaboratives in the UK. Lancet Gastroenterol Hepatol. 2017;2(4):247–8.
9. Weiser TG, Regenbogen SE, Thompson KD, Haynes AB, Lipsitz SR, Berry WR, et al. An estimation of the global volume of surgery: a modelling strategy based on available data. Lancet. 2008;372(9633):139–44.

10. McQueen KA, Coonan T, Derbew M, Tangi V, Bickler S, Banguti P, et al. The 2015 Bangkok Global Surgery Declaration: a call to the global health community to promote implementation of the world health assembly resolution for surgery and anaesthesia care. World J Surg. 2017;41(1):7–9.
11. Bhangu A, Kolias AG, Pinkney T, Hall NJ, Fitzgerald JE. Surgical research collaboratives in the UK. Lancet. 2013;382(9898):1091–2.
12. Pinkney TD, Calvert M, Bartlett DC, Gheorghe A, Redman V, Dowswell G, et al. Impact of wound edge protection devices on surgical site infection after laparotomy: multicentre randomised controlled trial (ROSSINI Trial). BMJ. 2013;347:f4305.
13. National Surgical Research C. Multicentre observational study of performance variation in provision and outcome of emergency appendicectomy. Br J Surg. 2013;100(9):1240–52.
14. Collaborators DT, West Midlands Research C. Dexamethasone versus standard treatment for postoperative nausea and vomiting in gastrointestinal surgery: randomised controlled trial (DREAMS Trial). Br Med J. 2017;357:j1455.
15. Reinforcement of Closure of Stoma Site C, the West Midlands Research C. Randomized controlled trial of standard closure of a stoma site vs biological mesh reinforcement: study protocol of the ROCSS trial. Color Dis. 2018;20(2):O46–54.
16. Chari A, Jamjoom AA, Edlmann E, Ahmed AI, Coulter IC, Ma R, et al. The British Neurosurgical Trainee Research Collaborative: five years on. Acta Neurochir. 2018;160(1):23–8.
17. Brennan PM, Kolias AG, Joannides AJ, Shapey J, Marcus HJ, Gregson BA, et al. The management and outcome for patients with chronic subdural hematoma: a prospective, multicenter, observational cohort study in the United Kingdom. J Neurosurg. 2017;127(4):732–9.
18. Collaborative ST. Impact of postoperative non-steroidal anti-inflammatory drugs on adverse events after gastrointestinal surgery. Br J Surg. 2014;101(11):1413–23.
19. Collaborative ST. Multicentre prospective cohort study of body mass index and postoperative complications following gastrointestinal surgery. Br J Surg. 2016;103(9):1157–72.
20. Collaborative ST. Outcomes after kidney injury in surgery (OAKS): protocol for a multicentre, observational cohort study of acute kidney injury following major gastrointestinal and liver surgery. BMJ Open. 2016;6(1):e009812.
21. Collaborative ST. Association between peri-operative angiotensin-converting enzyme inhibitors and angiotensin-2 receptor blockers and acute kidney injury in major elective non-cardiac surgery: a multicentre, prospective cohort study. Anaesthesia. 2018;73(10):1214–22.
22. Chapman SJ, EuroSurg C. Ileus Management International (IMAGINE): protocol for a multicentre, observational study of ileus after colorectal surgery. Color Dis. 2018;20(1):O17–25.
23. EuroSurg Collaborative. Safety and efficacy of non-steroidal anti-inflammatory drugs to reduce ileus after colorectal surgery. Br J Surg. 2019; https://doi.org/10.1002/bjs.11326.
24. Mendelow AD, Gregson BA, Fernandes HM, Murray GD, Teasdale GM, Hope DT, et al. Early surgery versus initial conservative treatment in patients with spontaneous supratentorial intracerebral haematomas in the International Surgical Trial in Intracerebral Haemorrhage (STICH): a randomised trial. Lancet. 2005;365(9457):387–97.
25. Nepogodiev D, Martin J, Biccard B, Makupe A, Bhangu A, National Institute for Health Research Global Health Research Unit on Global S. Global burden of postoperative death. Lancet. 2019;393(10170):401.
26. Haynes AB, Weiser TG, Berry WR, Lipsitz SR, Breizat AH, Dellinger EP, et al. A surgical safety checklist to reduce morbidity and mortality in a global population. N Engl J Med. 2009;360(5):491–9.
27. Collaborators C, Roberts I, Shakur H, Afolabi A, Brohi K, Coats T, et al. The importance of early treatment with tranexamic acid in bleeding trauma patients: an exploratory analysis of the CRASH-2 randomised controlled trial. Lancet. 2011;377(9771):1096–101.
28. Collaborators C, Shakur H, Roberts I, Bautista R, Caballero J, Coats T, et al. Effects of tranexamic acid on death, vascular occlusive events, and blood transfusion in trauma patients with significant haemorrhage (CRASH-2): a randomised, placebo-controlled trial. Lancet. 2010;376(9734):23–32.
29. Weiser TG, Makary MA, Haynes AB, Dziekan G, Berry WR, Gawande AA, et al. Standardised metrics for global surgical surveillance. Lancet. 2009;374(9695):1113–7.

30. GlobalSurg C. Mortality of emergency abdominal surgery in high-, middle- and low-income countries. Br J Surg. 2016;103(8):971–88.
31. GlobalSurg C. Surgical site infection after gastrointestinal surgery in high-income, middle-income, and low-income countries: a prospective, international, multicentre cohort study. Lancet Infect Dis. 2018;18(5):516–25.
32. Dowswell G, Bartlett DC, Futaba K, Whisker L, Pinkney TD, West Midlands Research Collaborative BUK. How to set up and manage a trainee-led research collaborative. BMC Med Educ. 2014;14:94.
33. Aiken AM, Haddow JB, Symons NR, Kaptanis S, Katz-Summercorn AC, Debnath D, et al. Use of antibiotic prophylaxis in elective inguinal hernia repair in adults in London and south-east England: a cross-sectional survey. Hernia. 2013;17(5):657–64.
34. Collaborative UKNSR. Multicentre observational study of adherence to Sepsis Six guidelines in emergency general surgery. Br J Surg. 2017;104(2):e165–e71.
35. CholeS Study Group WMRC. Population-based cohort study of outcomes following cholecystectomy for benign gallbladder diseases. Br J Surg. 2016;103(12):1704–15.
36. Pearce L, Newton K, Smith SR, Barrow P, Smith J, Hancock L, et al. Multicentre observational study of outcomes after drainage of acute perianal abscess. Br J Surg. 2016;103(8):1063–8.

Chapter 19
Inclusion of Patient-Reported Outcomes in Clinical Trials

Lindsey M. Zhang, Cord Sturgeon, Anthony D. Yang, and Ryan P. Merkow

19.1 Introduction

When evaluating the outcome of an operation, surgeons are traditionally focused on issues of morbidity and mortality. Although it is extremely important to know metrics such as the occurrence of a surgical site infection or death after surgery, measuring the patient perspective is also necessary for determining the success of an operation. A patient-reported outcome (PRO) is defined by the U.S. Food and Drug Administration as "any report of the status of a patient's health condition that comes

L. M. Zhang
Department of Surgery, University of Chicago Medical Center, Chicago, IL, USA

Division of Research and Optimal Patient Care, American College of Surgeons, Chicago, IL, USA

C. Sturgeon (✉)
Division of Surgical Oncology, Department of Surgery, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA
e-mail: Cord.Sturgeon@nm.org

A. D. Yang
Division of Surgical Oncology, Department of Surgery, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA

Surgical Outcomes and Quality Improvement Center, Northwestern University, Chicago, IL, USA

R. P. Merkow
Division of Research and Optimal Patient Care, American College of Surgeons, Chicago, IL, USA

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

Surgical Outcomes and Quality Improvement Center, Northwestern University, Chicago, IL, USA

T. M. Pawlik, J. A. Sosa (eds.), *Clinical Trials*, Success in Academic Surgery,
https://doi.org/10.1007/978-3-030-35488-6_19

directly from the patient, without interpretation of the patient's response by a clinician or anyone else" [1]. An example of a PRO might include a patient's characterization of postoperative fatigue or their physical function satisfaction after an operation, as both determinations are made by the patients themselves and cannot be directly measured or observed. Studies have shown that clinical or physical assessments are not always reflective of how the patient actually functions or feels, highlighting the importance of incorporating PROs into surgical practice [2].

In recent years, both the research community and U.S. governmental regulatory bodies shifted focus to incorporate the patient's perspective in quality assessments. In 2010, the Patient-Centered Outcomes Research Institute (PCORI) was established as a part of the Patient Protection and Affordable Care Act. PCORI was born from a concern that research efforts in medicine fail to capture what matters most to patients and their families, and the Institute has since funded hundreds of studies focused on improving patient-centered care [3]. A unique aspect of many of the funded projects is that they are clinical trials which have a focus on PROs as either an outcome or intervention. Incorporation of PROs in clinical trials is a growing trend and will be an essential skill for academic clinicians. In this chapter, we will review the application of PROS into clinical trials including methods of integration, strategies for PRO measurement, and guidelines for protocol development and publication.

19.2 Integrating PROs into a Clinical Trial Design

In the world of scientific inquiry, the randomized clinical trial reigns supreme, providing the highest quality of evidence for healthcare decision-making. Use of PROs in clinical trials has increased steadily since the early 2000s, and encouraging results from these studies have shown that clinical trials are well positioned to improve patient-centered outcomes [4]. PROs can add value to clinical trials in many ways. This section will discuss different methods of incorporating PROs into clinical trials (Fig. 19.1) as well as several examples from the medical literature that highlight the impact of measuring the patient perspective in medicine.

19.2.1 Patient-Reported Outcomes as the Outcome

Evaluating PROs as a primary or secondary outcome provides unique insight into the patient's perception of the impact of an intervention. Consider, for example, a clinical trial comparing Drug A to Drug B. Traditional outcomes might discover that Drug B has better disease-free survival compared to Drug A, suggesting that Drug B is the superior medication. However, imagine if evaluation of PROs as a secondary outcome revealed that patients report severely reduced physical function and worse gastrointestinal symptoms with Drug B, suggesting that, outside of a clinical trial, there may be low drug adherence. Knowledge that a drug might be tolerated

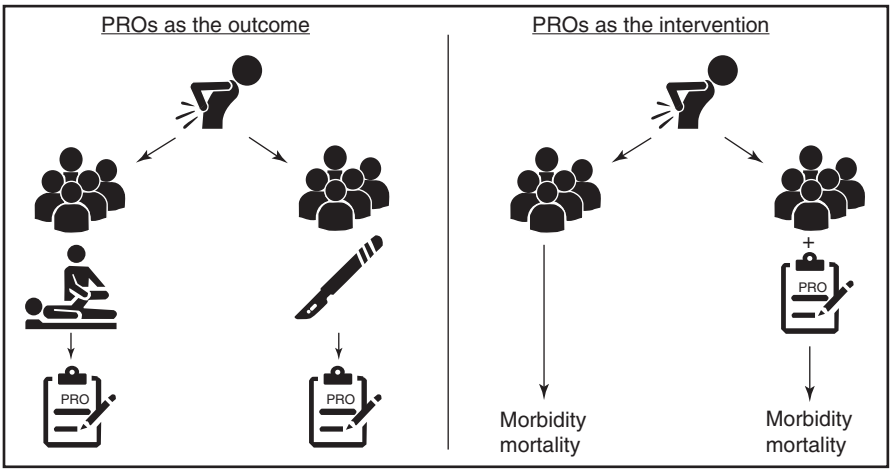


Fig. 19.1 Methods for incorporating PROs into clinical trials. Using a clinical trial for back pain as an example, participants can be assigned to receive physical therapy or surgery with scoring on a PRO measurement instrument evaluated as the outcome (Left). Alternatively, patients can be assigned to undergo routine PRO measurement and usual care as the intervention vs. usual care alone, and morbidity and mortality are evaluated as the outcome (Right)

poorly or frequently discontinued would be a key finding for this study. This example highlights the importance of measuring the patient's perspective when shaping patient-centered care. In addition to serving as a secondary outcome and providing context to other traditional health outcomes, PROs can also be measured as the primary outcome of a clinical trial.

In a study by van de Graaf et al., early intervention with arthroscopic partial meniscectomy was compared to physical therapy for knee function after a meniscal tear [5]. The primary outcome for this study was a patient-reported measure of knee function as evaluated by a scale called the International Knee Documentation Committee Subjective Knee Form. The results of the study showed that there was a similar increase in post-intervention patient reported knee function scores compared to pre-intervention in both the surgical and physical therapy arms. Currently, early surgery is a common treatment for meniscal tears, but the PROs of this study suggest that patient perceived outcomes of function are similar between both interventions. This finding allows clinicians to have a more patient-centered discussion regarding the two treatment options.

19.2.2 Patient-Reported Outcomes as the Intervention

In addition to serving as the outcome of a study, PROs can be incorporated into clinical trials as the intervention itself. Utilizing patient feedback as an intervention or part of an intervention can provide valuable knowledge on the effect of PROs on morbidity and mortality. In a 2017 study by Basch et al., the use of PROs during

chemotherapy for metastatic tumors was compared to usual care and the outcome of overall survival was evaluated [6]. The results of this study showed that the intervention arm, which randomized patients to have PROs integrated into the care pathway, had a median survival of 31 months, a statistically significant survival benefit compared to the usual care arm with a median survival of 26 months. Possible explanations for this finding include earlier symptom detection and more personalized care management, allowing for better chemotherapy tolerability and adherence.

A randomized clinical trial performed by Cleeland et al. evaluated not just the measurement of PROs but also the effect of PRO feedback to providers [7]. The focus of this study was on patients undergoing cancer surgery that included a thoracotomy, an operation with a notoriously difficult postoperative recovery. For all enrolled participants, PROs were elicited two times per week for a month in the post discharge phase of care. Those randomized into the intervention arm additionally had alerts sent to their physicians if PRO measurement reached a threshold score, suggesting the patient was experiencing severe symptoms. This was referred to as a "symptom threshold event." The control arm had no alerts to providers for symptom threshold events. Results of the study revealed that those in the intervention arm had a statistically significant decrease in the overall number of PRO measurements that reached the symptom threshold compared to the control arm. The results also showed a significantly faster decline in symptom threshold events over time for the intervention arm compared to the control. From this randomized clinical trial, the impact of evaluating PROs with internal feedback and clinician alerts demonstrates how profound the impact of PROs can be on patient care and comfort.

19.3 Measuring and Interpreting PROs

For researchers who want to incorporate PROs into clinical trials, as either the outcome or the intervention, it is critical to understand the appropriate measurement methods and tools. Measuring PROs in an objective manner is an inherently difficult task, as the purpose of a PRO is to assess a patient's subjective perception of his or her own health and function. Fortunately, there are several guidelines and examples of validated, objective ways to measure PROs that can be used to properly incorporate patient feedback into clinical trials.

In 2010, the Consensus-Based Standards for the Selection of Health Status Measurement Instruments, or the COSMIN study, published a checklist of necessary properties for a well-designed PRO measurement tool [8]. This checklist includes four major areas—Reliability, Validity, Responsiveness, and Interpretability. In short, a high-quality PRO measurement instrument should produce consistent results, and variance in outcome should be reflective of differences between patients rather than random error. The content and structure of the tool should adequately address the patient-reported measurement of interest and should be generalizable across ages and cultures. Responsiveness refers to an instrument's ability to detect change in answers over time, an especially important feature for any instrument

used to measure PROs in a clinical trial. Finally, a PRO measurement tool should produce results that can be interpreted and translated into clinically relevant information. The COSMIN checklist, updated in 2018 to the COSMIN Risk of Bias Checklist, can serve as a guideline for researchers interested in evaluating the quality of a PRO measurement tool or for developing a PRO instrument for use in their studies [9].

For researchers interested in using previously validated tools, several exist for measuring PROs. One example includes the Patient-Reported Outcome Measurement Information System, or PROMIS® tool, created in 2004 as an initiative of the National Institute of Health [10]. Within PROMIS, there are hundreds of different PRO measures spanning the domains of physical, social, and mental health for both adults and children (Fig. 19.2). The PROMIS measures are free and available for public use. Researchers can choose the PROs that they are interested in, such as depression or physical function, and access psychometrically validated measurement instruments for clinical or research use. When PROMIS measures are used, results are provided as a "*T*-score," where a *T*-score of 50 represents the reference score of the general U.S. population and every 10 points below or above 50 represent one standard deviation from the mean. Interpreting the results of a *T*-score

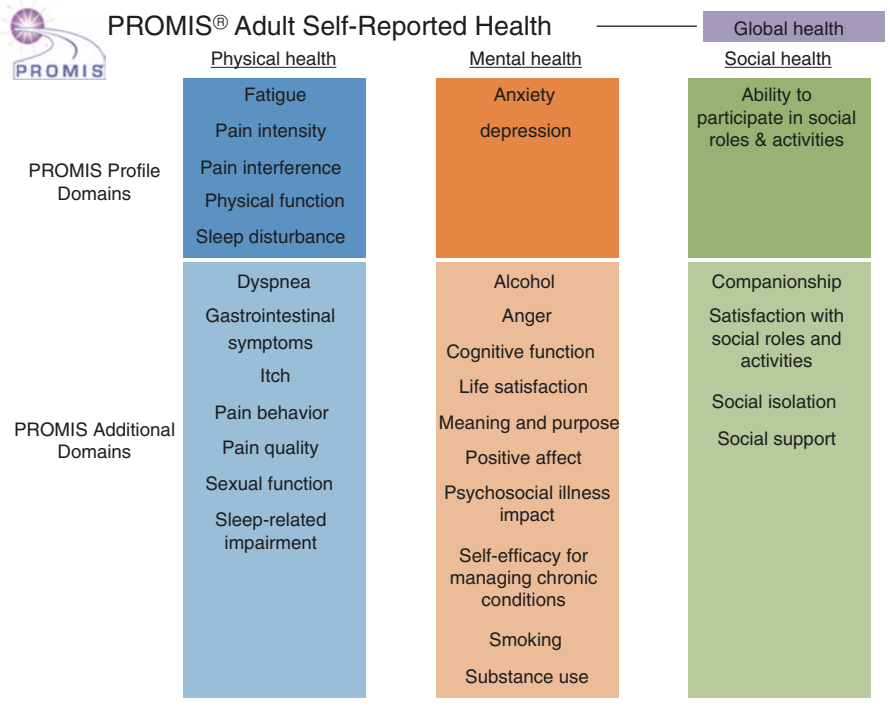


Fig. 19.2 Patient-reported outcome measurement information system (PROMIS) measure domains (© 2008–2019. Reprinted with permission, PROMIS Health Organization. PROMIS is a registered trademark of HHS)

depends on the PRO being measured. For example, using the PROMIS measure for depression, a *T*-score of 80 would mean more depressive symptoms compared to the general population. However, a *T*-score of 80 when using the PROMIS measure for physical function would mean better functionality compared to the population mean. PROMIS provides a wide variety of measurement options for interested researchers, but there are certainly other validated tests and tools available. For example, in the Basch study, PROs were measured using EQ-5D, a standardized instrument that measures patient mobility, ability for self-care and activity, pain, and anxiety [11]. The Cleeland study utilized the MD Anderson Symptom Inventory, a measurement instrument designed specifically for cancer patients to assess severity of symptoms and their impact on daily living [12]. Researchers should choose the instruments that best measure the PROS of interest and address the objectives of the study.

19.4 Guidelines for Including PROs in a Clinical Trial Protocol and Report

There are several guidelines available to researchers interested in including PROs in a clinical trial. The Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) statement provides a checklist of key topics to address in the development of a clinical trial protocol [13]. Aligned with the increased focus on PROs, the SPIRIT-PRO extension statement was published outlining the necessary aspects of including PROs in clinical trial protocols [14]. Additionally, the Consolidated Standards of Reporting Trials, or CONSORT group, which provides guidelines on how to write up and report the results of a clinical trial, has now published a CONSORT-PRO checklist [15, 16]. The inclusion of PROs into these consensus guideline statements highlights the burgeoning importance of the patient perspective for the future of clinical trials.

19.5 Limitations and Challenges to Using PROs in Clinical Trials

While there are numerous benefits to incorporating PROs into clinical trials, such an endeavor does not come without its challenges. Objectively measuring PROs requires a high-quality measurement instrument and, for some clinical studies, there may not be a previously validated tool for measuring the PROs of interest. Creating an appropriate test can be a difficult task for researchers to undertake and using an inappropriate tool may undermine the validity of the clinical trial. Furthermore, implementing these PRO instruments can be logistically challenging. In the Cleeland study, for example, PROs were measured twice weekly for a month for each patient.

While this provided an incredibly rich source of data for researchers to interpret, it likely required a significant investment of personnel, expertise, and time. Incorporating PROs into any clinical trial will demand extra resources, which may limit the feasibility of their integration. Additionally, surgeons interested in PROs in clinical trials must consider possible disruption to surgical workflow, challenges coordinating PRO measurement with electronic health records, and the potential need to risk adjust the inherently subjective information measured in PROs.

19.6 Conclusion

The field of surgery has changed immensely over the past centuries from what was once the purview of the town barber to what can truly be considered a well-researched and evidence-based area of medicine. As surgeons continue to learn more about the impact of both operative and non-operative treatments on patients, it has become strikingly clear that the next stage of evolution in surgery must be to incorporate the patient voice and experience as complementary outcomes to traditional measures. Although there remain challenges to including PROs in clinical trials, the incorporation of PROs provides the unique opportunity for clinicians to improve not only morbidity or mortality but also patients' functional outcome. As demonstrated by the examples detailed in this chapter, incorporating PROs into clinical trials can provide crucial information to help shape patient-centered care, and there is little doubt that PROs will continue to be an important, if not required, component of prospective studies in the future.

Acknowledgement R.P.M. is supported by the Agency for Research and Healthcare Quality (K12HS023011) and an Institutional Research Grant from the American Cancer Society (IRG-18-163-24).

References

1. FDA Guidance for Industry. Patient-reported outcome measures: use in medical product development to support labeling claims. 2009. http://www.fda.gov/downloads/Drugs/Guidances/UCM193282.pdf.
2. Sloman R, Rosen G, Rom M, Shir Y. Nurses assessment of pain in surgical patients. J Adv Nurs. 2005;52(2):125–32. https://doi.org/10.1111/j.1365-2648.2005.03573.x.
3. Selby JV, Beal AC, Frank L. The Patient-Centered Outcomes Research Institute (PCORI) national priorities for research and initial research agenda. JAMA. 2012;307(15):1583–4. https://doi.org/10.1001/jama.2012.500.
4. Mercieca-Bebber R, King MT, Calvert MJ, Stockler MR, Friedlander M. The importance of patient-reported outcomes in clinical trials and strategies for future optimization. Patient Relat Outcome Meas. 2018;9:353–67. https://doi.org/10.2147/prom.s156279.
5. van de Graaf VA, Noorduyn JCA, Willigenburg NW, et al. Effect of early surgery vs physical therapy on knee function among patients with nonobstructive meniscal tears: the

ESCAPE randomized clinical trial. JAMA. 2018;320(13):1328–37. https://doi.org/10.1001/jama.2018.13308.
6. Basch E, Deal AM, Dueck AC, et al. Overall survival results of a trial assessing patient-reported outcomes for symptom monitoring during routine cancer treatment. JAMA. 2017;318(2):197–8. https://doi.org/10.1001/jama.2017.7156.
7. Cleeland CS, Wang XS, Shi Q, et al. Automated symptom alerts reduce postoperative symptom severity after cancer surgery: a randomized controlled clinical trial. J Clin Oncol. 2011;29(8):994–1000. https://doi.org/10.1200/jco.2010.29.8315.
8. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res. 2010;19(4):539–49. https://doi.org/10.1007/s11136-010-9606-8.
9. Mokkink LB, Vet HCWD, Prinsen CAC, et al. COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. Qual Life Res. 2017;27(5):1171–9. https://doi.org/10.1007/s11136-017-1765-4.
10. Cella D, Yount S, Rothrock N, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. Med Care. 2007;45(5 Suppl 1):S3–S11.
11. EQ-5D. EuroQol Group. https://euroqol.org/.
12. MD Anderson Symptom Inventory (MDASI). MD Anderson Cancer Center. https://www.mdanderson.org/research/departments-labs-institutes/departments-divisions/symptom-research/symptom-assessment-tools/md-anderson-symptom-inventory.html.
13. Chan A, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, et al. Statement: defining standard protocol items for clinical trials. Ann Intern Med. 2013;158:200–7. https://doi.org/10.7326/0003-4819-158-3-201302050-00583.
14. Calvert M, Kyte D, Mercieca-Bebber R, et al. Guidelines for inclusion of patient-reported outcomes in clinical trial protocols: the SPIRIT-PRO extension. JAMA. 2018;319(5):483–94. https://doi.org/10.1001/jama.2017.21903.
15. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. Trials. 2010;3:c332. https://doi.org/10.1186/1745-6215-11-32.
16. Calvert M, Blazeby J, Altman DG, et al. Reporting of patient-reported outcomes in randomized trials: the CONSORT PRO extension. JAMA. 2013;309(8):814–22. https://doi.org/10.1001/jama.2013.879.

Chapter 20
Participation in Clinical Trials as a Clinical Trialist for the Community Surgeon

Jonah D. Klein and Ned Z. Carp

20.1 The History of Clinical Trials in Community Hospitals

In order to discuss the community surgeon's involvement in clinical trials, it is necessary to understand the evolution of how community clinicians have been involved in clinical trials. This evolution and transformation will originally focus on oncology. In the 1950s, the National Cancer Institute (NCI) began its Clinical Trials Cooperative Group Program, which has been the primary means of how community clinicians have been involved in national clinical trials [1]. Since its inception, the Cooperative Group Program has undergone consolidation and restructuring [2]. Over the first few decades of the cooperative group's existence, its primary objective was developing and testing new chemotherapeutic agents, but this transformed into various clinical trials for the treatment, prevention, and detection of disease. Several cooperative groups formed based on specialties (Fig. 20.1) which greatly enhanced the enrollment capabilities for clinical trials. In the more recent decades, most of the original cooperative groups restructured and underwent mergers to consolidate (Fig. 20.1). These mergers represent the present-day structure of the cooperative groups. The general organization of the current cooperative groups (ECOG-ACRIN, The Alliance, NRG, COG) include lead academic institutions, institutional affiliates, sub affiliates, and importantly community hospitals via the National Cancer Institute (NCI) Community Oncology Research Programs (NCORPs), further outlined in the next section.

Most community-based clinical trials involving cooperative groups have been in oncology. Since the 1970s, community hospitals have been involved in national research and clinical trials. At that time, it was evident that many patients sought care at community hospitals, but there was concern that community clinicians

J. D. Klein · N. Z. Carp (✉)
Department of Surgery, Lankenau Medical Center, Wynnewood, PA, USA
e-mail: carpn@mlhs.org

T. M. Pawlik, J. A. Sosa (eds.), *Clinical Trials*, Success in Academic Surgery,
https://doi.org/10.1007/978-3-030-35488-6_20

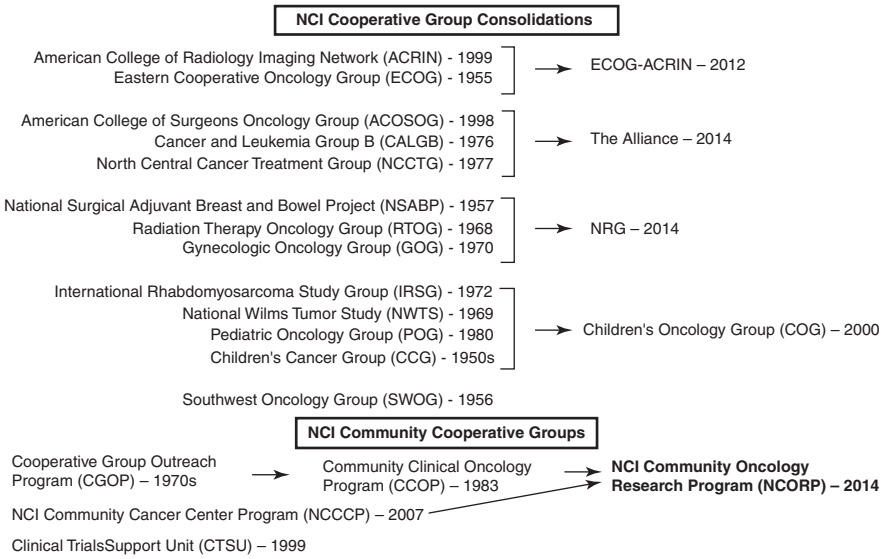


Fig. 20.1 Origination, mergers, and consolidations of oncology cooperative groups

wouldn't have interest in participating in clinical trials. There was further concern that community hospitals couldn't submit data of adequate quality, violate protocols, and not pay strict attention to eligibility criteria. To evaluate and mitigate these concerns, in 1976, the Eastern Cooperative Oncology Group (ECOG) initiated a program to involve community hospitals in multi-institutional clinical trials. At the time of this study initiative, there were 28 member institutions of ECOG, and after 5 years, ECOG expanded to 112 community hospitals, contributing an additional accrual of over 4500 patients in clinical trials. The data quality, eligibility standards, protocol compliance, and outcome measures were assessed, and they achieved equal outcomes when comparing community hospitals to ECOG member institutions [3].

The first cooperative group programs for community hospitals were the Cooperative Group Outreach Program (CGOP), and in 1983, the Community Clinical Oncology Program (CCOP) was initiated [4]. CCOP was structured to enable community hospitals to enroll patients in national clinical trials. There were 62 original participants in the CCOP, most of which had prior experience with the CGOP [4]. The Southwest Oncology Group (SOG) was one of the first cooperative groups to integrate these community research programs into their patient accrual increasing community physician participation [5].

The impact of community surgeons on clinical trials is demonstrated in 1985 when the *New England Journal of Medicine* published the five-year outcomes of the NSABP B-06 study. This study was the follow-up to the NSABP B-04 study that published six-year data in 1977 demonstrating that there is no difference in treatment failure or survival in clinically node negative patients undergoing the Halsted radical mastectomy, total mastectomy with radiation or total mastectomy followed by axillary dissection in those that develop positive nodes. The NSABP B-06 study

demonstrated that a segmental mastectomy followed by breast radiation, and adjuvant chemotherapy if there were positive nodes for stage I and II breast tumors, when compared to total mastectomy, did not demonstrate a survival difference [6, 7]. This study utilized patient participants from several CCOP institutions and other community hospital principal investigators, enabling a significant increase in the accrual of patients. This surgical clinical trial, inclusive of community hospitals, is the basis for the current standard of care of breast conservation surgery for invasive breast cancer. Another high-impact clinical trial published in 2010 was the NSABP B-32 trial demonstrating overall survival, disease-free survival and regional control was no different in patients with sentinel node negative invasive breast cancer, undergoing axillary surgery limited to sentinel node only versus axillary dissection. Again, a study with significant involvement of community surgeons, including lead authors, is the basis of most axillary surgery for breast cancer patients today [8].

In 1990, the NCI established the minority-based CCOP (MB-CCOP) to increase enrollment of racial and ethnic minorities and improve respective access to advances in cancer care. At the initiation of the CCOP, it was utilized for therapeutic trials, then expanded to cancer control and prevention trials. In 2007, the NCI established the NCI Community Cancer Center Program (NCCCP) aimed to enable community sites to support larger academic centers on trials involving prevention, screening, diagnosis, treatment, and end-of-life care [9]. In 2010, the Institute of Medicine (IOM) released a report to evaluate the cancer clinical trial structure and Cooperative Group Program design. In this report, they discuss the role of the community physician. It is noted that the majority of patients seek cancer treatment in the community setting, and of the enrolled patients in NCI's Cooperative Group Program, 65% of the patients were from community practices through CCOP and affiliates. By 2013, over 250,000 patients from the CCOP network were enrolled in clinical trials [10]. Despite the apparent successful enrollment, the structure of community involvement had many barriers including financial burdens, regulatory complexity, knowledge of trial availability, and attitudes toward community physician participation. Despite modest support grants available at the time, cost remained a barrier. The NCI recommended a consolidation and certification program to highlight community participants with attempts to overcome some of the barriers of community clinician participation [11]. In 2013, the IOM published a follow-up report highlighting the previously mentioned consolidations (Fig. 20.1) as well as the integration of the NCI CCOPs (NCCCP) and the MB-CCOP in the NCI Community Oncology Research Program (NCORP) [12].

20.2 National Cancer Institute Community Oncology Research Program (NCORP)

The NCI's NCORP program officially began in 2014 and is the present-day's community hospital's avenue for involvement in cancer clinical trials, as well as research on cancer disparities, prevention, screening, post-treatment management, cancer

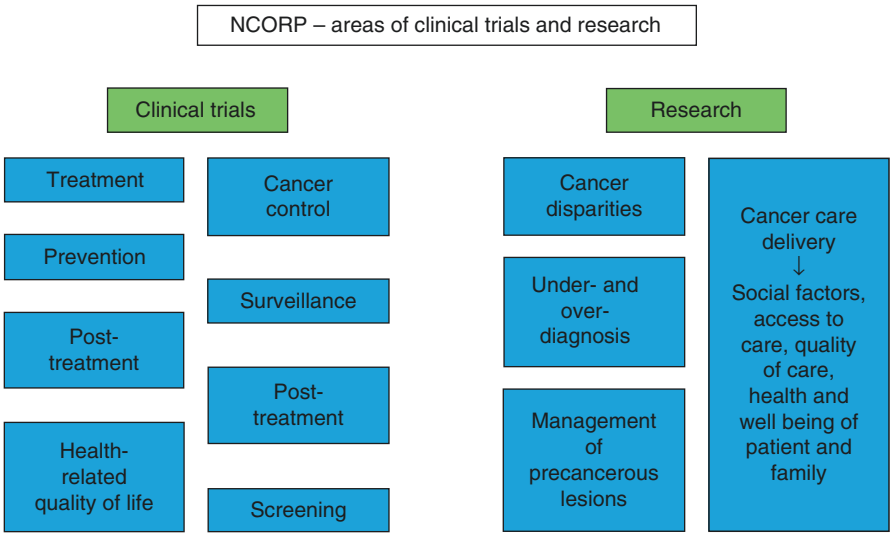


Fig. 20.2 Areas of research interests and requirements of the NCI Community Oncology Research Program (NCORP)

care delivery, and other fields (Fig. 20.2). At its inception, community institutions with National Clinical Trial Network (NCTN) research grants had their grants extended from 3 to 5 years. NCORP aimed at increasing accrual through web-based patient enrollment and structured rostering [12]. There were initially 53 new 5-year grants provided to researchers, broken down into the categories of 7 Research Bases, 34 Community Sites, and 12 Minority/Underserved Sites [13].

Research Bases served as the hubs for design and management of the clinical trials. These NCORP Research Bases are the previously mentioned consolidated cooperative groups (Alliance, COG, ECOR-ACRIN, NRG, SWOG) as well as the University of Rochester and Wake Forest. The Community Sites each have affiliated community healthcare centers and hospitals/practices and accrue patients into trials conducted by the Research Bases. For a list of current community sites and associated components and subcomponents, follow this link: https://ncorp.cancer.gov/findasite/. Included in this list are the Minority/Underserved Sites, who accrue patients in the same manner as the Community Sites, but must have at least a 30% resident ethnic and/or racial minorities [13]. Sites apply for grant funding through the NCTN, and the scientific leadership, statistics, and primary data management take place at the Research Bases.

The core principles of the NCORP include community organizations with various research capabilities, provide support to oncologic practices in different organizational settings, patient engagement, encourage commitment of support from participating organizations, and integrate disparities in cancer care [10]. Participating NCORP institutions are required to enroll patients in trials involving treatment, prevention, cancer control in addition to active research in cancer care

delivery disparities (Fig. 20.2). There is an encouragement for community sites to create a multisite collaborative effort to involve more community cancer patients in clinical trials.

In order for an NCORP site to find a clinical trial, they will go to the NCI NCORP website and then the "find a study" section: https://ncorp.cancer.gov/find-a-study/. In this section, the investigator will be able to determine the research base, determine the type of clinical trial (e.g., cancer prevention, treatment, cancer care delivery research, etc.), and then further be referenced to the formal clinical trial site. In order for an academic institution to apply to be a research base, or a community institution to apply as an NCORP site or Minority/Underserved Community Site, they should go to the applicants section of the NCI NCORP website: https://ncorp.cancer.gov/resources/applicants.html. Similarly, at smaller community centers, the NCORP does have the option to be a non-NCORP site and perform research in conjunction with an NCORP site, and applying for this type of participation is also available online.

An additional avenue for involvement in the NCORP, for sites that are not currently designated as NCORP institutions, is the Cancer Trials Support Unit (CTSU). This entity originally created in 1999 was designed to provide support and collaboration for clinical sites to participate in NCI-funded clinical trials. The CTSU has evolved and expanded to support many services in many different clinical trials and assists clinical sites with operational procedures and standardization. Getting started with the CTSU is an online process that can be initiated via this respective reference [14].

20.3 Resources Needed and Infrastructure for Successful Clinical Trial Involvement

The infrastructure needed to build a successful community clinical trial program is not different from an academic center including funding, research personnel, and collaborators. The difference is the available funding, presence of research expertise, trained research nurses and data managers, a structure for human subject research training, other interested colleagues, and trained contracts and legal personnel. Please refer to the chapter on "building a research team" for details, but this chapter focuses specifically on the infrastructure desired for successful participation and enrollment in surgical clinical trials for the community surgeon.

One element of a research team are the members, the other element is the expertise and training. Crucial members of a research team are data managers, clinical and research nurses, budget and legal experts, other participating clinicians in addition to the principal investigator (PI), members of the institutional review board (IRB), and others. Community hospitals are unique from academic centers, in that there is a huge variability of depth of research personnel. At a small rural community hospital, enrollment may be challenging, as the clinical nurse may take on the

role of the research nurse and data manager, and there may not be in-house budget, legal, and contract expertise. Other community institutions may be part of a larger health system, may even have their own research institute, where there are several data managers for different specialties, and in-house legal expertise. Fewer personnel in smaller centers often lead to lower patient recruitment, given available resources and available time of research staff. Smaller community centers may have the principal investigator as the only investigator, and that surgeon may have to get consent the patient him or herself and go through all aspects of a clinical trial with the patient. At larger community centers, there may be many involved clinicians in addition to the PI, and the designated research nurse can review trials in detail and consent patients. Regardless, the resources available to the community investigator are nearly universally sparse compared to university centers.

The other aspect of the research team is training. Training also comes in many forms, both learning about the clinical trials, navigating larger databases, assisting in trial recruitment, as well as regulatory, ethics, and compliance training. With university institutions, when hiring and training research personnel, that expertise already exists within the institution, but depending on the size of the community center, surgeons may have to train their own research staff. Regarding ethics and compliance training, university and large community centers may subscribe to a program, such as the Collaborative Institutional Training Initiative (CITI), that provides human subjects research training modules. Smaller, non-affiliated community centers may have to seek this type of training elsewhere.

Legal and budgeting expertise is crucial to the implementation of clinical trials. This is to ensure that centers that participate in trials do not lose money and practice safely. An expert must be present at the community center, or a contracted service, to ensure appropriate reimbursement, perform a Medicare coverage analysis, determine what is standard of care, prepare a trial budget and possibly meet with an auditing committee. A contracts expert must be able to prepare contracts that can define involved parties, publication rights, confidentiality rights, termination rights, sponsor liability, and indemnification (important with NIH versus device trials). If there are not expert experienced personnel that can provide legal, budgeting, and contract services, this must be obtained or contracted to a reputable company as this is key in the successful implementation of clinical trials for both institutional finances and patient safety.

20.4 Factors Associated with Community Surgeon Involvement in Clinical Trials and Patient Recruitment

Since the initiation of the community-based clinical trials, physicians that participated in these groups had variable patient enrollment, many of which enrolled zero patients in a given year [15]. There have been studies that have evaluating factors associated with involvement and recruitment from the standpoint of the patient, physician, and institution.

The attitude that a physician has toward the value of clinical trials and the logistical ease of enrolling patients had direct correlation with more actively enrolling physicians (based on a retrospective analysis of CCOP) [15]. This factor is especially important at a community hospital, as the involvement in a given research study is often voluntary from the standpoint of the clinician. Being a principal investigator on a study also was associated with better recruitment. Surgeons were noted to be less likely to accrue patients in a clinical trial compared to medical oncologists; older physicians are less likely to accrue than younger colleagues.

Additionally, factors that were associated with organizational context improve patient recruitment and physician enrollment. These contextual factors include staff support for consenting and enrolling patients, institutional incentives and acknowledgements for enrollment, and training opportunities for physicians to learn about trial participation and patient enrollment [16]. Also, institutions involved in more cancer control trials and quality of life were more likely to accrue patients in clinical trials. Interestingly, the value clinicians place on involvement in clinical trials is independent of the organizational context in which they work, but both factors influence recruitment. Practice location and type has played a variable role in patient recruitment.

Only 3–5% of eligible patients are enrolled in clinical trials. Factors associated with enrollment include availability of trial, knowledge of availability of trial, trial exclusions, age, race, and gender [17]. Younger, Caucasian males are more likely to be clinical trial participants. Major barrier for patient involvement, in addition to knowledge of availability, include patient commitment and champions. Commitment and champions from the standpoint of patients, as well as other entities including advocacy groups and professional societies, will also raise awareness and increase enrollment in community centers. Another factor is to mitigate patient's preexisting perspectives on clinical trials. This may be more relevant in the community setting where patients aren't expecting active clinical trials. Clinicians must mitigate issues of mistrust that continue to loom from such cases as the Tuskegee syphilis experiment and inform patients of governing bodies such as the Office for Human Research Protections and mandated review board approvals to ensure the safety and ethical nature of trials. Patients may also be under the impression that clinicians accrue patients to get kickbacks from sponsoring companies. Clinicians must assure patients that the purpose of clinical trials is to improve outcomes and enhance quality of life.

20.5 Community Center Versus Academic Training Center

Community centers come with a wide range of academics and physician postgraduate training. Some, as in the authors', have medical and surgical resident and fellowship programs, whereas other community centers have none. Regardless, research in all of these institutions is feasible whether through NCORP directly, being an NCORP affiliate, through drug and device trials, or independent clinical trials. The primary differences between a community center and a university-based academic center are the availability of funds, institutional backing, and resources. And

this trend continues with the range of community centers depending on the training offered and the nature of the associated health system.

At smaller, or private, community centers, finding the resources to get involved in clinical trials can be very challenging. At larger academic centers, and even larger community centers, there is designated research staff that facilitates a large portion of the logistics and conversations. The surgeon can mention the trial and spark a patient's interest, but as far as scheduling, consenting, billing, detailed eligibility requirements, and other follow-up tasks, this is often relegated to the research staff designated to clinical trials. At smaller community centers, this responsibility often falls on the surgeon or the clinical nurse, and in a busy practice, it can be burdensome.

This chapter's section is designed to address those surgeons who may be at a small community hospital or based out of a more rural private practice, but would like to be involved in clinical trials. While, this setting makes involvement much more challenging, it is still possible. Suggested ways to initiate involvement is to select a mentor at a larger community site who is already involved in clinical trials, and/or attend a national or international society meeting on the topic of interest for clinical trial involvement. This is where interested physicians can meet clinicians and discuss common interests and inquire about involvement. These conversations, while informal, are what lead to awareness of clinical trials, networking with device and drug companies, and creation of a network of interested surgeons. Once involved with one trial, a clinician will often be sought by other companies or trialists for involvement in additional trials. And as one trial ends, the trial topic gets morphed into another follow-up trial continuing involvement.

Barriers that will be more difficult to overcome will be in the areas of cost, patient recruitment, and lack of personnel. Addition costs may be accrued when initiating clinical trials to have documents reviewed (contracts and IRBs) by privately hired legal personnel. Also, patients often seek community hospitals as their health center, and private practices for expertise, but not because of active clinical trials, which makes the trial recruitment process more difficult. The staff that is part of the surgeons practice must be willing to make some sacrifices, such as extra time and responsibilities to support the clinical trials that are being performed, in addition to the existing clinical practice. The surgeon must also make sacrifices and have understanding family and friends given that research time is added to, and doesn't take the place of, clinical time. But despite these added barriers and responsibilities, there are many community and private practice surgeons that have broken through and contributed to patient care improvement by initiating and persisting in clinical trials involvement.

20.6 Doing Clinical Trials with Pharma and Device Companies

For a detailed, comprehensive outline of how the Food and Drug Administration (FDA) regulates pharmaceutical and device trials, please refer to Chap. 10. This current chapter's section will not focus on drug and device classification and the steps to approval, but more so the advantages, disadvantages, and recruitment of

drug and device trials from the community surgeon's perspective, and how the community surgeon can get involved in these trials.

As an example of a device trial, ACOSOG published a phase II trial (Z1072) in 2016 on the success of cryoablation therapy in the treatment of invasive breast cancer [18]. This was a mutli-institutional surgical device clinical trial with participants from university and community health centers. All cryoablation procedures were performed by the Visica 2™ Treatment System and sponsored by the manufacturing company Sanarus®. Cryoablation was performed followed by mandated surgical resection. This trial demonstrated a complete pathologic response to cryoablation therapy and concordance with detection in MRI for tumors less than 1 cm, and the follow-up trial without surgical resection is accruing (FROST trial). Community centers had leading recruitment numbers for this trial.

An example of a surgical drug trial was a study performed evaluating Alvimopan and its role in postoperative ileus in major abdominal surgery [19]. This was a randomized, placebo controlled, multi-institutional, both community and academic centers, in North America, evaluating the use of Alvimopan in a modified intent-to-treat study sponsored by the two involved pharmaceutical companies. This study demonstrated that Alvimopan use significantly decreased time to gastrointestinal recovery and hospital discharge in patients undergoing bowel resection or radical hysterectomy.

There are also trials in vascular surgery that extend to the community surgeon. An example of this is the currently recruiting TRANSCEND trial [20]. This is a randomized, single-blind, non-inferiority trial evaluating patients with symptomatic peripheral arterial disease of the femoropopliteal system to the SurVeil drug-coated balloon, versus the IN.PACT Admiral drug-coated balloon. This is a device sponsored trial and involves both large community and academic centers.

The primary advantage of a community surgeon's involvement in trials involving investigational new drugs or investigational devices is the respective company's sponsorship. Most of the time, trial participants, recruiting physicians, and insurance companies are not charged for new drugs, except allowing of sponsors to break even on the cost of the drug by insurance companies. Regarding investigational devices, the U.S. Centers for Medicare and Medicaid Services (CMS), in most cases, will reimburse the cost of the device up to the current day standard or similarly approved marketed product. The physician participating in the study must submit a Request for Reimbursement Letter to verify CMS reimbursement prior to conducting the study. For device trials, depending on the complexity and novelty of the device, there may be required physician and staff training [21].

20.7 Non-pharmaceutical, Non-device Community Clinical Trials

Let's start with another example of an impactful, purely surgical, clinical trial that had contributions by the community surgeon. The Multicenter Selective Lymphadenectomy Trials (MSLT-I and MLST-II) published in the *New England Journal of Medicine* in 2006 and 2017, respectively, evaluated the contribution of

the sentinel lymph node biopsy and completion lymph node dissection in patients with melanoma. In MSLT-I, patients with intermediate thickness of primary cutaneous melanoma were randomized to excision with post-operative observation of lymph nodes and lymphadenectomy, if clinical relapse or sentinel lymph node biopsy and immediate lymphadenectomy were positive. The 5-year survival rate of those patients in the sentinel lymph node group was 20% greater compared to the observation group [22]. The MSLT-II trial evaluated similar melanoma patients with sentinel lymph node metastasis, and compared serial ultrasound observation to immediate lymphadenectomy, and immediate lymphadenectomy did not demonstrate improved melanoma-specific survival [23]. Both of these trials included contributions from community centers.

Given the infrastructure that the NCI has in place for community involvement in clinical trials, most clinical trials performed to date have been in the field of oncology. There are previously mentioned device trials with endovascular devices, breast cancer, post-operative medications, but it is important to discuss clinical trials that have been performed in surgery, outside device and drug trials and outside of the field of oncology, and to further discuss the challenges involved with such trials.

A groundbreaking trial published in the *New England Journal of Medicine* in 1990 was the Program on the Surgical Control of Hyperlipidemia (POSCH) trial. This trial was a randomized clinical trial testing the partial ileal bypass operation on the mortality due to coronary heart disease. Eligible patients hospitalized after their first myocardial infarction were randomized to hospital discharge versus continued hospitalization for partial ileal bypass. In follow up, fewer patients in the surgery arm were on cholesterol lowering medications, there were lower total cholesterol and LDL cholesterol and higher HDL cholesterol in the surgery arm and a 22% risk reduction in overall mortality and 28% risk reduction in cardiovascular mortality in the surgery arm. Most importantly for this chapter, of the 838 patients from four institutions in the trial, 184 of them were from a community hospital, which was the second highest accruing center. While ileal bypass is now not routinely used due to the safety and efficacy of statins, this trial demonstrates an early surgical trial involving a highly accruing community center [24]. In 2010, the *Annals of Surgery* published a 25-year follow-up evaluation demonstrating an increase in cardiovascular mortality and all-cause mortality in the control arm [25].

One of the biggest challenges in a community surgeon getting involved in a clinical trial, such as the previously mentioned NSABP surgical trials, or the POSCH trial, is the standardization from one surgeon to another and from one institution to another. As in these cases, when the surgical intervention itself is the clinical trial, there are many elements that could potentially lead to challenges, biases, and questioned results such as anesthesia use, preoperative and postoperative care, different surgeons at different institutions, follow-up regimens, among others [26]. While this problem exists in the academic setting, a benefit of community hospital involvement in surgical trials is increased accrual, and to take advantage of this, often trials will involve multiple centers. There are varying degrees of standardization that will dictate the strength of a given trial's outcome. The surgical intervention itself can be standardized with respect to its written description, preparation, dissection,

resection, approach, closure, or a selection of these factors. The same goes for the other variable factors. The interventions need to be monitored for compliance to these standards, and there may be a training and preceptorship period prior to community surgeon involvement in order to optimize standardization.

20.8 Lifestyle and What's in It for the Busy Community Surgeon?

Major barriers in the utilization of the Cooperative Group Programs by community hospitals have been recruitment and reimbursement. Community physicians desire patient involvement, but with increasing patient load and decreasing time per patient, there often isn't time left for appropriate discussion of trial protocols. And if there is time made for protocol discussion, there are often additional tests and follow-up not often covered by insurers. There hasn't been enough grant support, and physicians fear increasing out-of-pockets expenses for patients. The time it takes for a "paying customer" is taken for a patient involved in a clinical trial. It is sometimes difficult to justify the busy community surgeon's involvement in clinical trials on top of existing clinical duties. These barriers do exist, but they are not inhibitory and can be overcome with collegiality, persistence, hard work, and desire [10].

With all these barriers, what's in it for the surgeon? This is an important question to ask. With financial strains, resource limitations, intense regulatory processes, immense time commitment, and a lack of financial incentive, why should a busy community surgeon take on clinical trials? The answer is simply in the community surgeon's desire to provide the best care to patients, optimize care for the best outcome, and satisfy one's academic curiosity. Given that the majority of cancer patients are seen and treated in community centers, it is imperative to continue to enroll these patients in clinical trials via active physician recruitment and participation. Community surgeon clinical trial involvement is crucial to ensure the best care, best outcomes and best quality of life for their patients, and contribute to global advances in health care to provide the best care, best outcomes, and best quality of life to all surgical patients.

Acknowledgements Paul B. Gilman, M.D., Albert DeNittis, M.D., M.S., Jarrod Kauffman, M.D., Alexander Uribe, M.D., John Wellenbach, John H. Marks, M.D.

References

1. National Center for Health Statistics, Center for Disease Control. 2017. https://www.cdc.gov/nchs/hus/contents2017.htm?search=Hospital_use. Accessed 26 Mar 2019.
2. Bauer TL Sr, Boughey JC. Community-based physicians and hospitals need to participate in clinical trials. Bull Am Coll Surg. 2014;99(4):46–7.

3. Begg CB, Carbone PP, Elson PJ, Zelen M. Participation of community hospitals in clinical trials—analysis of five years of experience in the eastern cooperative oncology group. N Engl J Med. 1982;306(18):1076–80.
4. Zon RT. Reforming the community research program: from community clinical oncology program to the National Cancer Institute Community Oncology Research Program. ASCO Educ Book. 2014;2014:e116–9.
5. Coltman CA Jr. Organizations and structures that facilitate community research. Semin Oncol. 1994;21(4 Suppl 7):107–11.
6. Fisher B, Bauer M, et al. Five-year results of a randomized clinical trial comparing total mastectomy and segmental mastectomy with or without radiation in the treatment of breast cancer. N Engl J Med. 1985;312(11):665–73.
7. Fisher B, Montague E, et al. Comparison of radical mastectomy with alternative treatments for primary breast cancer. Cancer. 1977;39(6):2827–39.
8. Krag DN, Anderson SJ, et al. Sentinel-lymph-node resection compared with conventional axillary-lymph-node dissection in clinically node-negative patients with breast cancer: overall survival findings from the NSABP B-32 randomised phase 3 trial. Lancet Oncol. 2010;11(10):927–33.
9. Siegel RD, et al. National collaborative to improve oncology practice: the National Cancer Institute Community Cancer Centers Program quality oncology practice initiative experience. J Oncol Pract. 2009;5(6):276–81.
10. McCaskill-Stevens W. The NCI Community Oncology Research Program: what every clinician needs to know. ASCO Educ Book. 2013;2013:e84–9.
11. Institute of Medicine. Participation by community physicians. In: Institute of Medicine, editor. A national cancer clinical trials system for the 21st century: reinvigorating the NCI Cooperative Group Program. Washington, DC: The National Academies Press; 2010. p. 197–9.
12. Institute of Medicine. Tapping community practices. In: Institute of Medicine, editor. Implementing a national cancer clinical trials system for the 21st century: second workshop summary. Washington, DC: The National Academies Press; 2013. p. 55–9.
13. NP Office. NCI Community Oncology Research Program (NCORP) Gets, NCI news note, 1. Bethesda, MD: NCI; 2014.
14. Cancer Trials Support Unit: a service of the National Cancer Institute. https://www.ctsu.org/Public/Default.aspx. Accessed 15 Mar 2019.
15. Jacobs SR, et al. Organizational and physician factors associated with patient enrollment in cancer clinical trials. Clin Trials. 2014;11(5):565–75.
16. Jacobs SR. Achieving high cancer control trial enrollment in the community setting: an analysis of the Community Clinical Oncology Program. Contemp Clin Trials. 2013;34(2):320–5.
17. SCM Cancer Network. Oncology. 2019. https://www.cancernetwork.com/article/inadequate-awareness-and-participation-cancer-clinical-trials-community-oncology-setting. Accessed 13 Mar 2019.
18. Simmonr RM. A phase II trial exploring the success of cryoablation therapy in the treatment of invasive breast carcinoma: results from ACOSOG (Alliance) Z1072. Ann Surg Oncol. 2016;23(8):2438–45.
19. Wolff BG. Alvimopan, a novel, peripherally acting μ opioid antagonist results of a multicenter, randomized, double-blind, placebo-controlled, phase III trial of major abdominal surgery and postoperative ileus. Ann Surg. 2004;240(4):728–35.
20. Safety and Efficacy of the SurVeil™ Drug-Coated Balloon (TRANSCEND). 2018. https://clinicaltrials.gov/ct2/show/NCT03241459. Accessed 20 Mar 2019.
21. Blake-Michaels M. Applied clinical trials. 2010. http://www.appliedclinicaltrialsonline.com/considerations-medical-device-trials. Accessed 13 Mar 2019.
22. Morton DL. Sentinel-node biopsy or nodal observation in melanoma. N Engl J Med. 2006;355(13):1307–17.
23. Faries MB. Completion dissection or observation for sentinel-node metastasis in melanoma. N Engl J Med. 2017;376(23):2211–22.

24. Buchwald H. Effect of partial ileal bypass surgery on mortality and morbidity from coronary heart disease in patients with hypercholesterolemia. Report of the Program on the Surgical Control of the Hyperlipidemias (POSCH). N Engl J Med. 1990;323(14):946–55.
25. Buchwald H. Overall mortality, incremental life expectancy, and cause of death at 25 years in the program on the surgical control of the hyperlipidemias. Ann Surg. 2010;251(6):1034–40.
26. Blencowe NS. Standardizing and monitoring the delivery of surgical interventions in randomized clinical trials. Br J Surg. 2016;103(10):1377–84.

Index

T. M. Pawlik, J. A. Sosa (eds.), *Clinical Trials*, Success in Academic Surgery,
https://doi.org/10.1007/978-3-030-35488-6