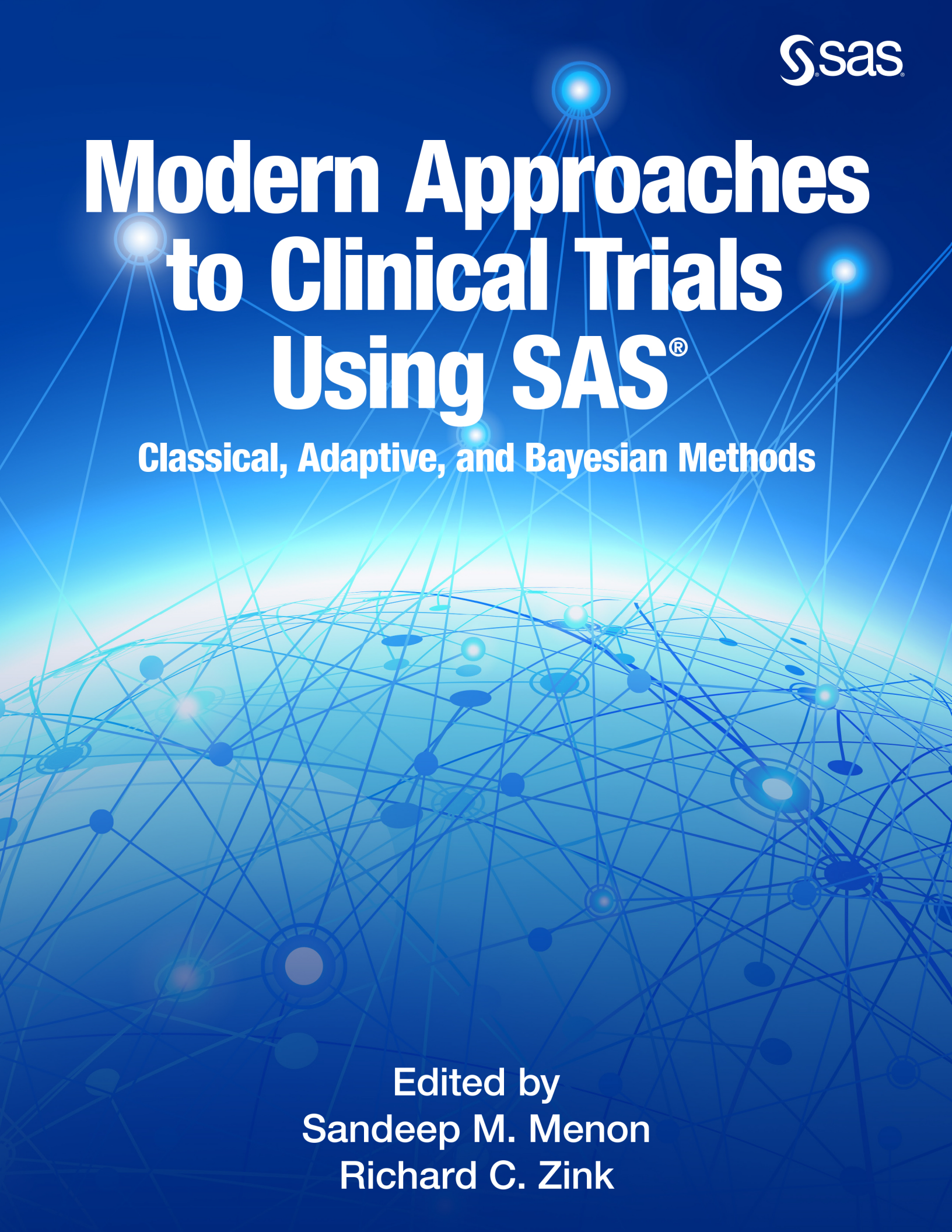
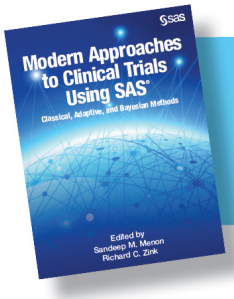
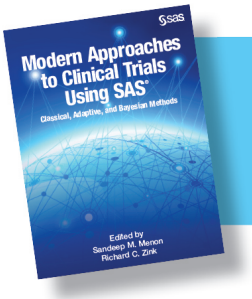

sas
Modern Approaches to Clinical Trials Using SAS®
Classical, Adaptive, and Bayesian Methods
Edited by
Sandeep M. Menon
Richard C. Zink



Contents



Chapter 1

Overview of Clinical Trials in Support of Drug Development

1.1 Introduction

Clinical testing of a drug to support its marketing authorization is often characterized by four phases. Here, we use the word "drug" broadly for a drug or a biologic. Three of the four phases are before the drug is marketed (pre-marketing) and one is afterwards (post-marketing). During the first phase (phase I), researchers investigate what the human body will do to a drug in terms of drug absorption, distribution, metabolism and excretion. The investigation is typically conducted in healthy human volunteers, except for cytotoxic drugs. For cytotoxic drugs, phase I is often conducted in patients with very few therapeutic options due to the anticipated toxicities and uncertainty about a drug's benefits. When a drug is designed to target a receptor or induce a certain biomarker response, phase I trials can sometimes investigate what the drug does to the body. Phase I investigation usually consists of single-dose and multi-dose escalations to understand the common adverse reactions of a drug and what would be the drug's dose-limiting toxicities. If the drug's safety profile is judged to be acceptable relative to its potential (and yet to be observed) benefit at this stage, the development will progress to the

second stage (phase II) with a recommended dose range. The number of volunteers included in phase I testing normally ranges between 20 and 80, but could be higher if phase I includes an assessment of the drug's mechanism of action or an early investigation of the drug's efficacy.

The second phase focuses on a drug's efficacy in patients with a targeted disorder. Clinical trials at this stage are also designed to determine dose(s), whose benefit-risk profile warrants further investigation in a confirmatory setting. Multiple doses within the dose range identified from phase I are typically studied during this phase. Occasionally, a sponsor may have to conduct more than one study if the doses chosen in the initial dose-response study are not adequate to estimate the dose-response relationship. This could occur if the doses selected initially are too high (e.g. near the plateau of the dose-response curve). To reduce the chance of having to repeat a dose-response study, it is generally recommended to include 4-7 doses in a wide dose range (the ratio of the maximum dose to the minimum dose ideally will be at least 10) in the dose-finding study. The analysis of a dose-finding study should focus on modeling the dose-response relationship instead of making pairwise comparisons between each dose and the control [1].

Phase II is typically the time when researchers first learn about the beneficial effect of a drug. It also has the highest attrition rate among the three pre-marketing phases. Therefore, if a drug is not a viable candidate, it is best to recognize this fact as soon as possible. This objective plus fewer regulatory requirements at this stage offer opportunities for out-of-the-box thinking. For example, some developers have divided phase II into two stages. The first stage tests the proof of concept (POC) of the drug, using a high dose (e.g., the maximum tolerated dose identified in phase I) to investigate a drug's efficacy. If the drug does not demonstrate a clinically meaningful efficacy compared to the control in the POC study, there will be no need to conduct a dose-response study. Otherwise, the drug will be further tested in a dose-ranging study. This two-step process is often referred to as phase IIa and phase IIb (see, for example, [2]). To streamline work that is required to initiate sites and obtain approvals from multiple institutional review boards, some sponsors combine POC and dose-response studies in one protocol with an unblinded interim analysis at the end of the POC stage. The sponsor will review results from the POC stage but use only data from the second stage to estimate the dose-response relationship. This strategy has the potential to reduce the so-called "white space" between phase IIa and phase IIb where the POC would be fully evaluated first and then the dose-response study would be planned.

Depending on the target disorders, phase II testing traditionally consists of 100-300 patients. Despite strong advocacy by researchers like [2] to use a modeling approach to analyzing dose-response data, some sponsors continue to rely on pairwise comparisons to design and analyze dose-response studies. There has been renewed emphasis that the selection of dose(s) is an estimation problem, and that this problem could be addressed more efficiently by using a modeling approach [3]. In addition, Pinheiro et al. have shown that even 300 patients in a dose-ranging study may not be enough to adequately identify the optimal dose based on a pre-set criterion [4].

If a drug meets the efficacy requirement and passes the initial benefit-risk assessment, it will be further tested to confirm its efficacy. This is the final stage of clinical testing before most drugs receive regulatory approval for marketing. This phase (phase III) enrolls a greater number of patients who are more heterogeneous in their demographic and baseline disease status. It is also at this stage that the majority of pre-marketing safety data are collected. Since a major objective of phase III is to confirm a drug's effect, analyses focus on testing pre-specified hypotheses with adequate control for the chance of making an erroneous claim of a positive drug effect. Operations at this stage require protecting a trial's integrity carefully so that trial results could be interpreted with confidence. The number of patients included at this stage typically ranges between 1,000 and 5,000. Drugs for orphan diseases will enroll much fewer patients while drugs that

are designed to reduce the risk of a clinical endpoint may require thousands, if not tens of thousands of patients. In addition, more patients will be needed if the drug is developed for multiple disorders simultaneously. An example for developing multiple indications simultaneously is antibiotics.

After a drug's effect is confirmed and benefit-risk assessment supports its use in the target population, the manufacturer of the drug will file a marketing application with regulatory agencies, typically in multiple countries. Nearly all applications are for the adult population initially. If the product is expected to be used in the pediatric population, a manufacturer will often have an ongoing pediatric development program or have a plan to initiate pediatric trials at the time of the initial marketing application. The marketing application may be for a single indication or for multiple indications. If the application is approved, the drug will be commercially available to the public. A manufacturer could choose to conduct additional studies to further test the drug in the indicated population(s), or in pediatric patients with the indicated disorder(s), or comparing the drug head-to-head with an approved drug for the same disorder(s), or for additional usages. Sometimes, a manufacturer conducts post-marketing studies to meet regulatory requirements as a condition for the marketing approval. This phase is often referred to as phase IV.

Another way to characterize the four phases of drug development is by the type of studies that are conducted during these 4 phases [5]. The types of studies conducted can be described as human pharmacology studies (phase I), therapeutic exploratory studies (phase II), therapeutic confirmatory studies (phase III), and therapeutic use studies (phase IV).

There are notable exceptions to the process described above. Many cancer drugs were initially granted accelerated approval based on tumor response rates observed in phase II trials. Some of the phase II trials may be single-arm studies. A condition for the accelerated approval is that the observed efficacy in phase II needs to be confirmed in randomized phase III trials. Depending on the type of cancer, the endpoint used in phase III trials can be progression-free survival or overall survival. When overall survival is not the primary endpoint in a phase III study, regulators often require that the new drug does not compromise overall survival. Drugs used to treat rare diseases could be approved based on phase II results also. The development pathway for each drug requires careful planning with input from regulatory agencies.

On 09 July 2012, the US Congress signed the Food and Drug Administration (FDA) Safety and Innovation Act. The Act allows the FDA to designate a drug as a breakthrough therapy if (1) the drug, used alone or in combination with other drugs, is intended to treat a serious or life-threatening disease or condition; and (2) preliminary clinical evidence indicates that the drug may demonstrate substantial improvement over existing therapies on at least one clinically significant endpoint. A manufacturer can submit the breakthrough designation request to the FDA for their drug and the agency has 60 days to grant or deny the request. Once a drug is designated as a breakthrough therapy, the FDA will expedite the development and review of such drug. The breakthrough designation can be withdrawn after granting [6].

Drug development has always been a high-risk enterprise. The success rate of developing an approved drug has decreased in recent years [7-9]. In 2004, the FDA in the United States (US) issued a Critical Path Initiative Document, in which the FDA quoted a "current" success rate of around 8% and a historical success rate of 14% [10]. To help lift the stagnation around drug development, the FDA encouraged innovations in many areas of drug discovery, development, and manufacturing. In the area of clinical development, the FDA encouraged, among several things, more efficient clinical trial designs. While looking for more efficient study designs has always been an area of intense research interest for many scientists, the need to look for new design options has accelerated since 2004. A class of designs beyond the traditional group sequential design

has emerged from these efforts. A common feature of these designs is to use interim data of a trial to modify certain aspects of the trial so that the trial can better address the questions it is designed to answer.

In "1.2 Evolution of Clinical Trials and the Emergence of Guidance Documents" on page 4 through "1.5 Widespread Research on Adaptive Designs Since the Turn of the 21st Century" on page 9, we discuss the evolution of clinical trials conducted to evaluate drugs. The evolution began with fixed trials, often done in a single or a few centers, to the more complex multi-center adaptive trials conducted by many manufacturers today. Group sequential design, which is an adaptive design, emerged in the early 70s. As the trial community began to embrace group sequential design in the 80s, researchers also began to develop designs using continual reassessment methods to search for the maximum tolerated dose in phase I cancer trials. Sample size re-estimation, both blinded and unblinded, was developed in the 90s and early part of the 21st century. During the first decade of the 21st century, significant efforts were dedicated to adaptive dose-ranging studies. Many of these designs are discussed in great detail in this book with companion SAS code to assist in their implementation.

As treatment became more personalized, adaptive designs have been proposed to help select the patient population for whom a new drug may be more effective. As better computational tools became more readily available, designs that incorporate information outside of the trial using Bayesian methodology have been explored and implemented. Despite the tremendous progress made over the past three decades, many challenges and opportunities in designing, conducting, and analyzing adaptive trials remain. We discuss some of them in "1.6 Opportunities and Challenges in Designing, Conducting, and Analyzing Adaptive Trials" on page 11.

We conclude this chapter with a discussion of the future adaptive trials to support drug development in "1.7 The Future of Adaptive Trials in Clinical Drug Development" on page 13.

1.2 Evolution of Clinical Trials and the Emergence of Guidance Documents

It took the pharmaceutical industry many years to reach the relatively mature state of drug development today. In 1962, the US Congress passed the Kefauver-Harris (KH) Amendment to the Federal Food, Drug, and Cosmetic Act of 1938 [11]. The amendment required drug manufacturers to prove the effectiveness and safety of their drugs in adequate and well-controlled investigations before receiving marketing approvals. Prior to the amendment, a manufacturer did not have to prove the effectiveness of a drug before marketing it.

It is not hard to imagine what drug manufacturers had to go through to comply with the KH Amendment initially. Thanks to the large polio vaccine trials in the 50s and 60s, the medical community was generally aware of the importance to randomize trial subjects in order to assess the effect of a new treatment against a comparator when the Amendment took effect. Still, the early randomized and controlled trials conducted by manufacturers were relatively simple and often took place in a single center or a few centers. It was not unusual for investigators to analyze data collected at their sites at that time. This practice began to change as drug companies began to employ statisticians in the mid 60s. Industry statisticians were initially hired to develop randomization codes and analyze data. It took several years for industry statisticians to get involved in designing drug trials. All early industry-sponsored trials used fixed designs, meaning that once a trial was started, the trial would continue until the planned number of patients was enrolled. While a trial could be stopped for safety reasons, there was no chance to stop the trial

early for efficacy, for futility, or to make modifications to the trial based on unblinded interim results. The concept of a pre-specified statistical analysis plan, signed off prior to database lock, did not exist.

While drug companies took steps to develop infrastructure for adequate and well-controlled trials, the National Institutes of Health (NIH) in the US led the way in increasing the standards for the design and conduct of clinical trials. In the 60s and 70s, the National Heart Institute within the NIH launched several ambitious projects to understand and manage an individual's risk for cardiovascular events. Randomized trials launched for this goal were typically large and required enrollment at multiple sites for the trials to complete within a reasonable time period. This practical need began the era of multi-center trials. Besides recruiting at a faster pace, multi-center trials allowed trial findings to generalize more broadly to the target population because trial results came from many investigators.

Even though the NIH provided oversight to these early multi-center cardiovascular trials sponsored by the Institute, statistical leadership at the NIH realized the need for a more organized way to monitor such trials and to potentially terminate the trials early for non-safety-related reasons. For example, it would be unethical to continue a trial if interim data clearly demonstrated one treatment was much better than the other. The same statistical leaders also recognized that by looking at trial data regularly and allowing the trial to stop early to declare efficacy, one could inflate the overall type I error rate. The above thinking led to the formation of a committee to formally review, at regular intervals, accumulating data on safety, efficacy, and trial conduct. The proposed committee is the forefather of the data monitoring committee (DMC) as it is known today [12]. The experiences led to the Greenberg Report in 1967, which was subsequently published in 1988 [13]. The Greenberg Report discusses the organization, review, and administration of cooperative studies. Another document of historical importance is the report from the Coronary Drug Project Research Group on the practical aspects of decision making in clinical trials [14]. The need to control the overall type I error rate due to multiple testing of the same hypothesis motivated statistical researchers at the NIH and elsewhere to initiate research on methods to control the type I error rate in the presence of interim efficacy analyses.

Pharmaceutical companies began testing cardiovascular drugs and cancer regimens in the late 70s. Following the NIH model, drug companies recruited patients from multiple centers. It did not take long for multi-center trials to become the standard for clinical trials to evaluate drugs in other therapeutic areas also. Furthermore, it was a common practice by the 90s to have a DMC for an industry-sponsored trial with mortality or serious morbidity as the primary endpoint.

Many regulatory guidance documents were issued in the 80s and 90s. For example, the Committee for Proprietary Medicinal Products (CPMP) in Europe issued a guidance entitled "Biostatistical Methodology in Clinical Trials in Applications for Marketing Authorisations for Medicinal Products" (December, 1994). The Japanese Ministry of Health and Welfare issued "Guidelines on the Statistical Analysis of Clinical Studies" (March, 1992). The US FDA issued a guidance entitled "Guideline for the Format and Content of the Clinical and Statistical Sections of a New Drug Application" (July, 1988). To help harmonize the technical requirements for registration of pharmaceuticals for human use worldwide, regulators and representatives from the pharmaceutical industry in Europe, Japan, and the US jointly developed common scientific and technical aspects of drug registration at the beginning of the 90s. The collaboration led to the formation of the International Conference on Harmonisation (ICH) and the publication of many guidance documents on quality, safety, and efficacy pertaining to drug registration. ICH issued a guidance document on statistical principles for clinical trials (ICH E9) for adoption in all ICH regions in 1998 [15]. ICH E9 drew from the respective guidance documents in the three regions mentioned above.

At the time that ICH E9 was issued, group sequential design was the most commonly applied design that included an interim analysis. ICH E9 acknowledges that changes in inclusion and exclusion criteria may result from medical knowledge external of the trial or from interim analyses of the ongoing trial. However, E9 states that changes should be made without breaking the blind and should always be described by a protocol amendment that covers any statistical consequences arising from the changes. E9 also acknowledges the potential need to check the assumptions underlying the original sample size calculation and adjust the sample size if necessary. However, the discussion on sample size adjustment in E9 pertains to blinded sample size adjustment that does not require unblinding treatment information for individual patients.

In 2007, the Committee for Medicinal Products for Human Use (CHMP, previously the CPMP) of the European Medicines Agency published a reflection paper on adaptive designs for confirmatory trials [16]. In 2010, the US FDA issued its own draft guidance on adaptive designs [17]. Both guidances caution about operational bias and adaptation-induced type I error inflation for confirmatory trials. The US draft guidance places adaptive designs into two categories: generally "well-understood" and "less well-understood" designs. "Less well-understood" adaptive designs include dose-selection adaptation, sample size re-estimation based on observed treatment effect, population or endpoint adaptation based on observed treatment effect, adaptation of multiple design features in one study, among others. It has been more than five years since the publication of the draft guidance and much knowledge has been gained on designs originally classified as "less well-understood." As experience accumulates, we expect some of the "less well-understood" designs will become "well-understood".

1.3 Emergence of Group Sequential Designs in the 70s and 80s

While the theory of group sequential design dates back to 1969, actual application began in the 1970s [18,19]. Canner notes the early evolution of applying multiplicity-adjusted analyses along with an external monitoring board in the Coronary Drug Project (CDP) [20]. For the first two years of CDP, investigators were informed of interim data by treatment group. Subsequently, perhaps the first external data and safety monitoring committee (DSMC) was formed to be the only reviewers of data summary by treatment group for the remainder of the trial. This trial also had what we now might call an executive committee (termed the CDP Policy Board then) that was charged with acting on DSMC recommendations. While formal stopping rules were not in place, there was an awareness of multiplicity issues associated with multiple active treatment groups and analyses at multiple time points, which may have resulted in an overall type I error rate on the order of 30% to 35%, if nominal Z-value cutoff for a two-sided significance level of 0.05 had been used repeatedly.

DeMets, Furberg and Friedman note that the Greenberg Report ensured that all cooperative group studies funded by the National Heart Institute and its successors had a separate monitoring committee to review interim results [21, p5]. A commonly cited example is the BHAT trial that began in 1978 and employed an O'Brien-Fleming boundary for group sequential monitoring of efficacy every 6 months [19]. The trial was stopped in 1981 after the O'Brien-Fleming efficacy boundary was crossed at an interim analysis.

Several papers summarize the early data-monitoring practice at one of the National Cancer Institute's cooperative groups, the Southwest Oncology Group (SWOG) [22,23]. They note that prior to 1984, unblinded interim results were routinely shared with study investigators and often published. The philosophy at the time was that those responsible

for the study should also be involved in the interim evaluations of safety and efficacy. Cancer researchers felt that the model of independent DMCs used in other NIH institutes was not feasible in trials conducted by the cancer cooperative groups [22]. There were noted examples where interim results were later reversed and situations where studies could not be completed due to the public sharing of interim results. As a result, starting in 1985, SWOG established a formal DMC. While toxicity was still shared with investigators in an unblinded fashion, formal group sequential stopping rules for efficacy were implemented using either Haybittle-Peto or O'Brien-Fleming bounds [24-26]. Interim efficacy results were reviewed by the DMC only.

Jennison and Turnbull provide a brief history of the theory and methods for sequential and group sequential designs, including citations for more complete histories [27, pp 5-11]. They note the work of Pocock as a key motivator for the use of group sequential designs by providing "clear guidelines for the implementation of group sequential designs attaining type I error and power requirements" [28]. The commonly used O'Brien and Fleming stopping rules came shortly thereafter, followed by developments that allow more flexible timing of interim analyses, such as the spending function methods of Lan and DeMets [26,29]. Pampallona and Tsiatis use boundary families to allow early stopping based on futility in demonstrating superiority of a new therapy over a standard [30]. Pampallona, Tsiatis and Kim extend the work of Pampallona and Tsiatis [31].

The 90s also saw aggressive pursuits of drugs to treat patients with the human immunodeficiency virus (HIV). The urgency in developing promising medicines provided a strong incentive for early monitoring of HIV trials for efficacy. This was supported by the cooperative groups and pharmaceutical industry, which was engaged in HIV trials, by patient advocacy groups, and by regulators at the FDA [32]. Finkelstein notes, for example, that the AIDS Clinical Trial Group trial #981 initiated in 1989 applied a one-sided group sequential boundary based on the Lan-DeMets spending function approximation to an O'Brien-Fleming design [33].

One of the authors of this chapter worked at Centocor in the 90s. We share two Centocor development programs as an example to illustrate the move to group sequential design by an industry sponsor. The example highlights the potential perils of inadequate documentation related to interim monitoring and benefits of group sequential design [32]. Both programs were to develop monoclonal antibodies to treat conditions that had irreversible consequences for patients. The conditions had few treatment options and, therefore, represented an urgent unmet medical need. As such, studies that investigated new treatment options merited interim monitoring to determine when study objectives had been achieved or if risk was excessive. In a first pivotal trial for one program, FDA reviewers felt that the company had not adequately documented that an interim change in the statistical analysis plan was made without incorporating information from unblinded interim results and, therefore, asked the company to perform a second pivotal trial. The second pivotal trial was unsuccessful when excess mortality was demonstrated at its first interim analysis. In a subsequent program, group sequential designs were incorporated into trials studying the effect of abciximab (a potent platelet inhibitor) to prevent acute ischemic events in patients undergoing coronary interventions. Three trials (EPIC, EPILOG, and CAPTURE) were conducted in the abciximab program. Both EPIC and EPILOG compared two abciximab-containing treatments to a standard therapy while CAPTURE was a two-arm trial [34-36]. The treatment regimens studied, particularly in the first trial (EPIC), had the potential for both substantial efficacy and substantial risk and thus merited interim monitoring for both safety and efficacy. EPIC proceeded past interim analyses and demonstrated efficacy at the final analysis. EPILOG and CAPTURE were stopped early due to demonstrated efficacy at interim analyses. These trials were all performed as industry collaborations with academic research organizations who were experienced in randomized clinical trials. All trials used independent external DMCs. Innovations to accommodate comparisons of multiple experimental arms were

achieved with modifications of the freely available FORTRAN programs from the University of Wisconsin [37].

Many statisticians found career opportunities in the pharmaceutical industry in the 90s. The influx of statisticians to the industry greatly expanded in-house statistical support to clinical trials. Statisticians' presence and the establishment of ICH helped increase the rigor of industry-sponsored clinical trials. In addition to contributing to the design, conduct, analysis, and interpretation of clinical trials, pharmaceutical statisticians also engaged in methodology research to help make the drug development process more efficient.

Group sequential designs are covered extensively in Chapter 2.

1.4 Emergence of Adaptive Designs in the 90s

Sequential and group sequential designs are a special kind of adaptive designs. While group sequential designs originated in the 60s, one can probably credit Bauer's work as the origin of what some refer to today as adaptive design [38]. Bauer first described sample size adaptation based on results of an unblinded interim analysis [38]. Bauer and coauthors gave a historical overview of the history of confirmatory adaptive designs over the 25 years since 1989 [39]. They describe the early days of adaptive design research, review the key methodological concepts, and summarize regulatory and industry perspectives on adaptive designs. The overview includes an extensive list of references (178 of them) and discusses the concepts of conditional power, conditional error and combination tests as the cornerstones for many approaches. It concludes with a critical review of how expectations from the beginning of the adaptive design journey were fulfilled, and it discusses potential reasons why the expectations were not fulfilled in some cases. Another good reference for adaptive designs is the book edited by He, Pinheiro and Kuznetsova [40].

Major reasons for adaptations include: (1) adapting sample size because of uncertainties concerning design parameters (variability, background rate, treatment effect) at the planning stage; (2) choosing among multiple possible treatments; and (3) adapting to a subpopulation where study treatment is the most effective. Choosing among treatments includes selecting doses in dose-finding studies and selecting among different treatment regimens. Both types of treatment selection are covered in this book.

Initial research on adaptive designs focused heavily on modifying sample size of a clinical trial. Sample size adaptation is discussed in great detail in Chapter 3. Early methods use conditional power [38,41,42]. These approaches led to discussions regarding design efficiency, which in turn led to improvements such as the promising zone design [43,44]. While there are other techniques for sample size adaptation, not all of them have received the same level of software support as have methods based on conditional error/combination tests. Examples include optimized sample size adaptation methods [45-48].

One approach is to use information-based group sequential design to adapt sample size [49]. Wan and coauthors suggest a relatively efficient sample size adaptation allowing only one alternative sample size in order to limit potential reverse engineering that could produce an estimate for the interim treatment effect [48,50]. This strategy is implemented using the promising zone code of Chapter 3 by setting the conditional power needed to adapt very high and setting an appropriate maximum sample size.

Another class of adaptive methods that emerged early focuses on response-based adaptive randomization. Response-based adaptive randomizations such as play-the-winner or randomized play-the-winner were proposed as early as the 60s and 70s

[51,52]. This class of adaptive randomization is discussed in Chapters 9 and 10. As noted by the authors for those two chapters, response-based adaptive randomization can be particularly valuable in studies where patients have a high risk of significant short-term outcomes, allowing a study to focus on the most effective treatments. While Chapter 10 focuses on binary outcomes, it references the broader applications of adaptive randomizations in the monograph by Hu and Rosenberger [53].

1.5 Widespread Research on Adaptive Designs Since the Turn of the 21st Century

1.5.1 Early Phase Oncology Designs

For many oncology development programs, the first clinical trials in humans are in cancer patients with a primary objective to estimate the maximum tolerated dose (MTD). A review paper by Le Tourneau et al. provides an overview of dose escalation methods for phase I oncology trials [54]. A 3+3 design has traditionally been used and continues to be used to estimate the MTD by some sponsors. A 3+3 design tests 3 patients at a dose initially. If none of the 3 patients has what is referred to as a dose limiting toxicity (DLT), the next higher dose will be studied. If 2 or more out of 3 patients have a DLT, the dose is considered toxic and will be excluded from further consideration. If 1 out of the first 3 patients at a dose has a DLT, another 3 patients will be enrolled at the same dose. If no more patients among the new cohort have the DLT, the dose is considered tolerable and the study can escalate to the next dose. Otherwise, the dose is considered intolerable and will be excluded. Once an intolerable dose is identified, if the dose below it has only been studied in 3 patients, another 3 will be given the same dose. If more than 1 patient has a DLT, then the dose is considered toxic and excluded. The maximum tolerated dose is the highest dose studied that was not discontinued per the algorithm above. Once an MTD is determined, some trials employing the 3+3 design will enroll additional patients (e.g. 12 or 24) at the MTD to investigate early signs of efficacy. Dose-escalation under the 3+3 design algorithm, while safe, often escalates through doses slowly and could be ineffective in finding an MTD.

One popular approach that has been proposed to improve upon the algorithm-based 3+3 design is the continual reassessment method (CRM) [55]. This approach is covered in Chapter 5 with further developments and SAS programs to support implementation. CRM is a Bayesian dose-finding method that adapts the up-and-down dose selection during a trial based on a modeled dose-toxicity curve. Another alternative to the 3+3 design is the modified toxicity profile interval proposed by Ji at al. [56]. The latter makes dose adjustments based on a table that can be generated at the beginning of the study according to a specified target DLT rate. The dose adjustment decisions have been implemented in Excel by Ji and coauthors.

1.5.2 Multiplicity in Adaptive Designs

Multiplicity arises frequently in multi-stage trials when conclusions may be based on interim data. For confirmatory trials, it is important to strongly control the overall type I error rate over multiple hypotheses tested or the number of times a hypothesis is tested. While solutions to some of these problems appeared in the 80s and 90s [27,Ch 15 and 16], a simple way to consider this for group sequential trials is to use a generalization of graphical methods for strong type I error control [57]. The graphical approach has also been extended to adaptive group sequential designs in Sugitani, Bretz, and Maurer [58].

Multiplicity also arises when researchers attempt to identify a subpopulation that experiences a better response (or experiences less side effects) to a treatment. Subpopulations could be defined by disease state at baseline or by a proteomic or genetic biomarker. Chapter 11 offers an extensive literature review on population enrichment designs and discusses enrichment strategies from a frequentist, Bayesian or a frequentist-Bayesian hybrid perspective.

1.5.3 Formation of the Adaptive Design Working Group

The intense interest in adaptive designs during the first decade of the 21st century motivated the formation of an Adaptive Designs Working Group (ADWG) in the spring of 2005 [59]. This was a collaboration that included contributions from industry, academia and regulatory authorities. Other than the group sequential design, adaptive design was still a relatively new concept for many drug companies at that time. Operational support such as randomization and drug supply management to support adaptive trials was not available in many organizations then. Furthermore, regulatory acceptance of the new adaptive designs was generally unknown. The objectives of the ADWG were to foster and facilitate wider usage and regulatory acceptance of properly designed and executed adaptive trials to support product development through a fact-based evaluation of the benefits and challenges associated with these designs [60]. The Group was initially sponsored by the Pharmaceutical Research and Manufacturers of America (PhRMA). In order to address the many aspects related to the design and implementation of adaptive trials, ADWG initiated many workstreams to kick off a broad range of activities. The activities included sponsoring workshops, giving short courses, and publishing research and consensus papers. A workstream on regulatory interactions reached out to regulators to discuss best adaptive design practice and share experience from implementing such designs [61]. A seminal white paper on best practice for adaptive trials was published by the Group in 2009 [62]. Workstreams that completed their objectives were sunset. New workstreams were initiated to tackle emerging issues.

The sponsorship for ADWG was officially transitioned from PhRMA to the Drug Information Association (DIA) in 2010. The name of the group was changed to the Adaptive Design Scientific Working Group (ADSWG) with expanded membership.

Because new investigators continue to join the clinical trial community, there is always a need to offer education and training. A long-running education and training activity of the Group is a monthly key opinion leader lecture series. The lecture series is free to all who are interested in adaptive designs. Early lectures focused on the theory underlying adaptive designs. Over time, the lectures expanded to practice and lessons learned from implementation. Some lectures focused on adaptive trials that were used to support regulatory submissions. A recurring theme is the importance of thorough upfront planning required of adaptive trials. The lecture series was still ongoing in October 2015 when this chapter went into printing.

1.5.4 Opportunities in the Learning Phase

An equally influential working group formed about the same time as the ADWG was the Adaptive Dose Ranging Studies Working Group (ADRS WG), again under the auspices of PhRMA. ADRS WG focused on the quantitative evaluation of adaptive designs and model-based methods for estimating dose-response relationships. A major objective of ADRS WG was to recommend when adaptive dose-ranging studies could be used and how much benefit they could be expected to bring. A series of white papers was published by the ADRS WG including [4, 63]. Major recommendations from the Group include the need to place dose selection in the broader context of the overall development program, and not restrict it to only the phase IIB stage. In addition, the WG recommends evaluating the impact of the choice of dose-ranging design and analysis on

the probability of success (PoS) of phase III and, ultimately, the expected net present value of a drug candidate. The ADRS WG was merged with the DIA ADSWG in early 2010. The work by the ADRS WG and continuing work by researchers on dose-response studies reminds researchers of the many opportunities to improve on how we design and analyze dose-response studies.

Thomas et al. analyzed dose-response studies conducted by a large pharmaceutical company for small molecules over a 10-year period (1998-2009) [64]. They also examined dose-response studies conducted by other drug companies [65]. They concluded that the dosing range and the number of doses tested were generally inadequate to characterize the dose-response relationship appropriately. They found that more than half of the studies they examined had a dose range (maximum dose divided by the minimum dose) less than 20. In many cases, lower doses were omitted from the original studies, causing the need for additional dose-response studies before phase III or a marketed dose to be lowered after product launch. Thomas et al. consider dose ranges less than 20-fold dubious to estimate parameters of the $E_{\max}$ model, the dose-response curve most commonly observed to fit the data. A dose range close to 100-fold would be more appropriate, in their opinion.

Dose-response is a critical stage in drug development. Getting the dose right at this stage critically impacts the chance of success in the confirmatory stage. Some simple adaptations at this stage could be useful [66]. For example, trialists could add a lower dose or a higher dose after an interim analysis. They could add a dose that is between two doses already included in the study to better estimate the sharpest part of a dose-response curve. These types of simple adaptations could help us better estimate the dose-response curve and select a dose or doses for phase III trials if the development program moves into the confirmatory phase.

More general dose-finding designs for studies outside of oncology are considered in Chapter 6 from the classical dose-finding perspective. Chapters 7 and 8 cover flexible modeling approaches.

1.5.5 Software

Since the objective of this book is to provide information as well as implementation of design and analysis of clinical trials using SAS, many SAS programs are included in this book. One other important SAS reference for adaptive design has been recently updated [67]. Chang provides some guidance on available SAS software for adaptive design (e.g., **`seqdesign`** and **`seqtest`**) as well as providing macros for many other types of adaptive designs. A good summary of other available adaptive design software can be found in Tymofyeyev [68].

1.6 Opportunities and Challenges in Designing, Conducting, and Analyzing Adaptive Trials

1.6.1 Logistics in Trial Execution

Implementing an adaptive trial requires operational support on many fronts. These include a versatile randomization system, nimble drug supply management, data monitoring support, timely access to fit-for-use data, process and documentation control. A detailed document that describes the rationale and execution of the pre-planned adaptations should be prepared in advance. The document should also describe the interim analysis plan including who will conduct the interim analysis and whether an

internal or external DMC will review the interim data. For confirmatory trials, access to interim results by the sponsor should be strictly controlled. If sponsor access is foreseen as a possibility for unique circumstances, the interim plan should describe in detail who within the sponsor may have access and under what circumstances. The plan should describe steps taken to prevent additional access of information within the sponsor organization. If sponsor access did occur, it should be clearly documented in the clinical study report. In the latter case, the clinical study report should discuss steps taken to minimize potential bias that may be introduced to the trial as a result of sponsor's access to interim data.

In general, interim analysis should be conducted by a statistician independent of the study team, or by a statistician who is a member of the DMC for the study. The detailed document mentioned above should be endorsed by the DMC. When recommendations requiring actions for a confirmatory study arise, final decisions on actions are often made by the trial's executive (or steering) committee that normally includes sponsor representatives. Antonijevic et al. recommend that a single DMC with all necessary expertise be convened to monitor and execute the pre-planned adaptation(s) [69]. In other words, they advise against having a separate monitoring committee whose sole responsibility is to look after the implementation of pre-planned adaptations.

One key challenge in implementing adaptive trials is the management of drug supply if adaptions result in drug supply change. A member from the drug supply organization should be part of the clinical team that plans an adaptive trial. Manufacturing of clinical study supplies, especially when they need to be matched or blinded, is costly and can be resource intensive. In addition to upfront costs, issues such as expiry dates for existing drug must be considered. Therefore, it is important to get drug supply personnel involved in the planning of an adaptive trial as early as possible. The extent that drug supply and associated allocation (randomization) procedures can accommodate design changes is an important factor when considering an adaptive design.

Adaptive design requires timely access to fit-for-use data for interim decisions. If time-to-event endpoint requires adjudication, adjudication on the number of events required for the interim analysis needs to take place prior to the planned interim analysis. Failure in timely access to fit-for-use data could compromise the effectiveness of an adaptation plan. However, we want to point out that the need for fit-to-use data for interim analysis for an adaptive trial is the same as that needed for a trial using a traditional group sequential design.

1.6.2 Open Research Questions

There are many opportunities for continuous improvement to increase the value and acceptance of adaptive design in the next two decades. We cite 4 as examples below.

1. Statisticians have generally solved the problem of controlling the overall type I error rate for confirmatory adaptive trials although debates continue on whether simulations can sufficiently demonstrate type I error control. The challenge of estimating treatment effect in the presence of adaptations, on the other hand, remains a research question for some designs such as the promising zone design.
2. Statisticians should help set the right mindset about adaptive designs through education and training. There are many occasions when a classic fixed design is the best choice for a situation. Statisticians should work with internal and external decision makers to help lead the discussion on what are appropriate design options for a particular situation.
3. Statisticians need to be aware of the operational support needed to implement an adaptive trial. If the design is too complicated or if the infrastructure needed to support its implementation is not available, a team needs to consider carefully

whether an adaptive trial is a good choice even if the planned adaptations can result in some benefit.

4. Statisticians need to remember the assumptions underlying adaptive designs. For example, the assumption of a stationary population is critical for an adaptive trial. If some of the fundamental assumptions are in question, one needs to think hard about whether an adaptive design is appropriate for a situation.

1.7 The Future of Adaptive Trials in Clinical Drug Development

A recent study released by the Tuft's Center for the Study of Drug Development suggests that the average pre-tax industry cost to bring a new medicine to market is now around $2.56 billion USD [70]. The study included 106 investigational new drugs from 10 mid- to large-size pharmaceutical companies and the drugs were first tested in humans during 1995-2007. Cost included clinical development up to 2013. By comparison, in 2003, the cost was about $1.04 billion in the 2013 dollars.

DiMasi stated that the higher cost comes from clinical trials that are larger and more complex. In our opinion, some of the higher costs also resulted from increased regulations [70]. For example, since the FDA issued a guidance document on evaluating cardiovascular risk in new antidiabetic therapies to treat Type 2 diabetes mellitus (T2DM) in December 2008, all new drugs for T2DM approved since 2008 have been or are being evaluated in cardiovascular outcome trials [71]. In the context of these large cardiovascular trials for T2DM, Chapter 4 utilizes Bayesian methodologies to take advantage of historical information to improve trial efficiency. These outcome trials enroll thousands, if not tens of thousands, of diabetic patients who are at an increased risk for cardiovascular events. Despite the new requirement, manufacturers continue to pursue anti-diabetes drugs because as Gregg et al. predicted, lifetime risk of diagnosed diabetes from age 20 years onward is about 40%, nearly doubling the risk of those born a decade or so earlier [72]. Another factor contributing to the higher cost is a higher failure rate during the clinical development phase in recent years. For example, there has been no drug approved for Alzheimer's disease (AD) in the US since 2003 when memantine was last approved [73]. Decades of investment in AD drugs by many companies, either as symptomatic cognitive enhancing or disease-modifying agents, have failed to produce a single new approved product or a new product close to be approved as of 2015.

While drug developers have always been aware of the high risk associated with drug development, the substantial increase in development cost has begun to change the business operating models for the industry. In recent years, there has been a strategic move towards co-development between pharmaceutical companies, or to share financial burdens for product development between the private and public sectors. The high cost has also motivated many companies to look for data-driven quantitative approaches to make better decisions so that if a development program is to fail, it can be terminated earlier and more efficiently. This new direction will further increase interest in adaptive trials.

Against the backdrop of high cost for large development programs, there is also a shift for industry to invest more heavily in precision medicine or medicines for rare diseases. In 2014, the FDA approved 41 new molecular entities for marketing. According to Gulfo of Breakthrough Medical Innovations, 40% of the new molecular entities approved are for rare diseases, underscoring the industry's shift to niche products [74]. There are several reasons for this shift. First, a large number of products that are highly effective in treating common disorders (e.g. high cholesterol, high blood pressure, CNS disorders)

have become generic in the past decade. It is hard to demonstrate extra value beyond what the now generic products could offer. On the other hand, value proposition is easier for rare diseases, which have received less attention in the past and continue to represent unmet medical needs. The smaller patient populations afflicted with rare diseases require out-of-box thinking and nimble applications of innovative approaches to clinical development. Carefully planned adaptive design serves this need nicely.

Another emerging trend is the use of a platform (umbrella) trial to screen multiple product candidates in a single trial. This is in contrast to traditional trials that typically investigate a new treatment (with multiple doses in some cases) in a generally homogeneous population. A platform trial could be used to investigate a product in patients with different genotypes or phenotypes (enriched subpopulations), or could also be used to investigate different treatments in one population. A more sophisticated platform trial could study multiple treatments in multiple enriched patient subpopulations. Interim analyses are conducted in platform trials to decide if a particular treatment (together with a subpopulation in some cases) could be graduated from the trial and further investigated in a confirmatory setting. Alternatively, a treatment could be dropped from the trial and a new treatment added to the trial, continuing the trial beyond the original set of treatments. Platform trials that allow the introduction of new treatments are sometimes called perpetual trials for this reason. Platform trials are also called "master protocol" trials because one protocol governs the testing of many drugs.

A well-known platform trial in the oncology area is the I-SPY 2 trial [75]. This is a phase II neoadjuvant trial for women with large primary cancers of the breast. Breast tumor is characterized by its response to three receptors (estrogen, progesterone, and HER2), resulting in 8 tumor signatures. The trial investigates multiple regimens that include investigational products from pharmaceutical companies. The primary endpoint is pathologic complete response at 6 months after treatment initiation. Within each tumor signature, adaptive randomization to regimens is employed. The trial may graduate or terminate a regimen according to a pre-specified rule based on an interim Bayesian prediction of phase III success probability for a (regimen, signature) combination. If the regimen remains in the trial after the interim decision, assignment to that regimen will continue but be capped at a pre-specified maximum number. One major advantage of a trial like I-SPY 2 is the ability to learn during the trial on what regimen benefits which patient subpopulation, and learn this by borrowing information from other (regimen, signature) combinations.

Another example is the lung-MAP (lung master protocol) trial, a multi-arm, biomarker-driven clinical trial for patients with advanced squamous cell lung cancer that was initiated in June 2014 [76]. Lung-MAP is a public-private collaboration. Lung-MAP plans to initially test five experimental drugs—four targeted therapies and an anti-PD-L1 immunotherapy. Patients will be screened for over 200 cancer-related genes for genomic alterations. The results of the test will be used to assign each patient to the trial arm that is best matched to their tumor's genomic profile. The study can test up to 5-7 drugs at one time and can be amended to test additional new drugs as current drugs exit the trial. In addition to the efficiency gain from using the same control group from the design aspect, a trial using a master protocol takes advantage of existing infrastructure and patient outreach efforts on the operational side.

The potential value of platform trials is not limited to the oncology area. Across the globe, the pace of development of new antibiotic products has slowed noticeably from its peak in the 80s, creating a public health crisis with the rapid development of drug-resistant bacteria. In the US, the President's Council of Advisors on Science and Technology (2014) published a Report to the President on Combating Antibiotic Resistance in September 2014 [77]. The Report offers practical recommendations to the US Federal government for strengthening the US's ability to combat the rise in antibiotic-resistant bacteria. In the area of clinical trials to test new antibiotics, the Report recommends increasing trial efficiency through improved infrastructure and

focusing on patient populations with the most urgent need. On ways to make clinical testing more efficient, the Council suggested the formation of a robust, standing national clinical trials network for antibiotic testing. The recommended action plan for the network includes the development of platform trials for antibiotics, where multiple new agents from different sponsors can be evaluated concurrently. As the utility of platform trials for screening product candidates becomes better understood, we are likely to see more such trials in the future.

Other adaptive designs that have received increasing attention by researchers include sequential multiple assignment randomized trials (SMARTs) and sequential parallel comparison designs (SPCDs) [78, 79]. Both classes of designs involve additional randomizations, based on patients' response following the initial randomization. SPCDs are also viewed as a way to address high placebo response rates in trials involving the central nervous system. SMARTs were originally proposed to increase efficiency in behavior science trials and to investigate optimal treatment strategies, but the concept can be applied to other types of trials as well.

Elsäßer et al. examined 59 scientific advice letters given by the Scientific Advice Working Party (SAWP) of the CHMP that addressed adaptive study designs in phase II and phase III clinical trials between 01 Jan 2007 and 08 May 2012 [80]. According to the authors, the most frequently proposed adaptation was sample size re-estimation, followed by dropping of treatment arms, and population enrichment. Among the 59 proposals, 15 were accepted (25%) and 32 were conditionally accepted (54%) by CHMP/SAWP. Elsäßer and coauthors concluded that despite critical comments in some cases, a majority of the proposed adaptive clinical trials received an overall positive opinion. Among the 41 more recent cases (out of 59) with more information in the advice letters, CHMP/SAWP noted insufficient justifications of the proposed adaptations, type I error rate control, and bias in treatment effect estimate as the most frequent concerns.

By the end of the first decade in the 21st century, there was a wide-spread interest in adaptive trials in the pharmaceutical industry. In an article, Burman and Chuang-Stein asked whether the interest in adaptive designs was a short-lived fascination or a reflection that adaptive designs could become part of the future of clinical research [81]. Since 2009, the clinical trial research community has made tremendous progress in understanding when adaptive trials add value and when they do not. Some hard lessons were learned in the process. We have seen products approved using evidence from pivotal adaptive trials. Some of the adaptive features such as futility analysis and some form of sample size re-estimation have become routine features of many registration trials.

Based on our own experience, we can confidently predict that properly designed and carefully executed adaptive trials that are not overly complicated and fit well in the context of a development program will have a firm place in the clinical research. They will become important tools in our trial design armamentarium as we continue to look for more nimble and efficient strategies to develop new and valued products.

References

1. Pinheiro J. (2014). Session 2 Summary – Designs & Methods. Presentation at the European Medicines Agency/European Federation of Pharmaceutical Industries and Associations workshop on the importance of dose finding and dose selection for the successful development, licensing and lifecycle management of medicinal products. Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Presentation/

2015/01/WC500179787.pdf (http://www.ema.europa.eu/docs/en_GB/document_library/Presentation/2015/01/WC500179787.pdf).

2. Sheiner LB. (1997). Learning versus confirming in clinical drug development. *Clinical Pharmacology & Therapeutics* 61: 275-291.

3. European Medicines Agency. (2014). European Medicines Agency/European Federation of Pharmaceutical Industries and Associations workshop on the importance of dose finding and dose selection for the successful development, licensing and lifecycle management of medicinal products. Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Report/2015/04/WC500185864.pdf (http://www.ema.europa.eu/docs/en_GB/document_library/Report/2015/04/WC500185864.pdf).

4. Pinheiro J, Sax F, Antonijevic Z, Bornkamp B, Bretz F, Chuang-Stein C, Dragalin V, Fardipour P, Gallo P, Gillespie W, Hsu C-H, Miller F, Padmanabhan SK, Patel N, Perevozskaya I, Roy A, Sanil A & Smith JR. (2010). Adaptive and model-based dose-ranging trials: Quantitative evaluation and recommendations (with discussion). *Statistics in Biopharmaceutical Research* 2: 435-454.

5. International Conference on Harmonisation. (1997). E8: General considerations for clinical trials. Available at: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E8/Step4/E8_Guideline.pdf (http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E8/Step4/E8_Guideline.pdf).

6. US Food and Drug Administration. (2014). Fact Sheet: Breakthrough Therapies. Available at: http://www.fda.gov/regulatoryinformation/legislation/significantamendmentstothefdcact/fdasia/ucm329491.htm (http://www.fda.gov/regulatoryinformation/legislation/significantamendmentstothefdcact/fdasia/ucm329491.htm) (accessed 9 Oct. 2015).

7. DiMasi JA, Hansen RW & Grabowski HG. (2003). The price of innovation: new estimates of drug development costs. *Journal of Health Economics* 22: 151-85.

8. DiMasi JA, Feldman L, Seckler A & Wilson A. (2010). Trends in risks associated with new drug development: success rates for investigational drugs. *Clinical Pharmacology & Therapeutics* 87: 272-277.

9. DiMasi JA, Reichert M, Feldman L & Malins A. (2013). Clinical approval success rates for investigational cancer drugs. *Clinical Pharmacology & Therapeutics* 94: 329-335

10. US Food and Drug Administration. (2004). Innovation or stagnation: Challenge and opportunity on the critical path to new medical products. Available at: http://www.fda.gov/ScienceResearch/SpecialTopics/CriticalPathInitiative/CriticalPathOpportunitiesReports/ucm077262.htm#fig3 (http://www.fda.gov/ScienceResearch/SpecialTopics/CriticalPathInitiative/CriticalPathOpportunitiesReports/ucm077262.htm#fig3).

11. Krantz JC. (1966). New drugs and the Kefauver-Harris amendment. *Journal of New Drugs* 6: 77-79.

12. Ellenberg S, Fleming T & DeMets D. (2002). *Data Monitoring Committees in Clinical Trials: A Practical Perspective*. Chichester, England: John Wiley & Sons.

13. Greenberg Report. (1988). Organization, review, and administration of cooperative studies. *Controlled Clinical Trials* 9: 137–148.

14. Coronary Drug Project Research Group. (1981). Practical aspects of decision making in clinical trials: the Coronary Drug Project as a case study. *Control Clinical Trials* 1: 363-376.

15. International Conference on Harmonisation. (1998). E9: Statistical principles for clinical trials. Available at: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/Step4/E9_Guideline.pdf (http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/Step4/E9_Guideline.pdf).

16. Committee for Medicinal Products for Human Use. (2007). Reflection paper on methodological issues in confirmatory clinical trials planned with an adaptive design. Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003616.pdf (http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003616.pdf).

17. US Food and Drug Administration. (2010). Draft guidance for industry: Adaptive design clinical trials for drugs and biologics. Available at: http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM201790.pdf (http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM201790.pdf).

18. Armitage P, McPherson CK & Rowe BC. (1969). Repeated significance tests on accumulating data. *Journal of the Royal Statistics Society, Series A (Statistics in Society)* 132: 235-244.

19. DeMets DL, Hardy R, Friedman LM & Lan KKG. (1984). Statistical aspects of early termination in the Beta-Blocker Heart Attack Trial. *Controlled Clinical Trials* 5: 362-372.

20. Canner PL. (2006). Breaking new ground: Data monitoring in the Coronary Drug Project. In: DeMets DL, Furberg CD & Friedman LM, eds. *Data Monitoring in Clinical Trials: A Case Studies Approach*. New York: Springer.

21. DeMets DL, Furberg CD & Friedman LM. (2006). *Data Monitoring in Clinical Trials: A Case Studies Approach*. New York, New York: Springer.

22. Green S & Crowley J. (1993). Data monitoring committees for Southwest Oncology Group clinical trials. *Statistics in Medicine* 12: 451-455.

23. Crowley J, Green S, Liu PY & Wolf M. (1994). Data monitoring committees and early stopping guidelines. *Statistics in Medicine* 13: 1391-1399.

24. Haybittle JL. (1971). Repeated assessments of results in clinical trials of cancer treatment. *British Journal of Radiology* 44: 793-797.

25. Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, Mantel N, McPherson K, Peto J & Smith PG. (1976). Design and analysis of randomized clinical trials requiring prolonged observation of each patient: I. Introduction and design. *British Journal of Cancer* 34: 585-612.

26. O'Brien PC & Fleming TR. (1979). A multiple testing procedure for clinical trials. *Biometrika* 35: 549-556.

27. Jennison C & Turnbull BW. (2000). *Group Sequential Methods with Applications to Clinical Trials*. Boca Raton, Florida: Chapman and Hall/CRC.

28. Pocock SJ. (1977). Group sequential methods in the design and analysis of clinical trials. *Biometrika* 64: 191–199.

29. Lan KKG & DeMets DL. (1983). Discrete sequential boundaries for clinical trials. *Biometrika* 70: 659-663.

30. Pampallona S & Tsiatis AA. (1994). Group sequential designs for one-sided and two-sided hypothesis testing with provision for early stopping in favor of the null hypothesis. *Journal of Statistical Planning and Inference* 42: 19-35.

31. Pampallona S, Tsiatis AA & Kim KM. (2001). Interim monitoring of group sequential trials using spending functions for the Type I and Type II error probabilities. *Drug Information Journal* 35: 113-121.

32. Ellenberg SS & Siegel JP. (2006). FDA and clinical trial data monitoring committees. In: DeMets DL, Furberg CD & Friedman LM, eds. *Data Monitoring in Clinical Trials: A Case Studies Approach*. New York: Springer.

33. Finkelstein D. (2006). Data monitoring in the AIDS clinical trials group study #981: Conflicting interim results. In: DeMets DL, Furberg CD & Friedman LM, eds. *Data Monitoring in Clinical Trials: A Case Studies Approach*. New York: Springer.

34. EPIC Investigators. (1994). Use of a monoclonal antibody directed against the platelet glycoprotein IIb/IIIa receptor in high-risk coronary angioplasty. *New England Journal of Medicine* 330: 956-961.

35. EPILOG Investigators. (1997). Platelet glycoprotein IIb/IIIa receptor blockade and low-dose heparin during percutaneous coronary revascularization. *New England Journal of Medicine* 336: 1689-1696.

36. CAPTURE Investigators. (1997). Randomised placebo-controlled trial of abciximab before and during coronary intervention in refractory unstable angina: the CAPTURE study. *Lancet* 349: 1429-1435.

37. Reboussin DM, DeMets DL, Kim KM, Lan KKG. (2000). Computations for group sequential boundaries using the Lan-DeMets spending function method. *Controlled Clinical Trials* 21: 190–207.

38. Bauer P. (1989). Multistage testing with adaptive designs. *Biometrie und Informatik in Medizin und Biologie* 20: 130–148.

39. Bauer P, Bretz F, Dragalin V, König F & Wassmer G. (2015). Twenty-five years of confirmatory adaptive designs: opportunities and pitfalls. *Statistics in Medicine*. Available at: http://onlinelibrary.wiley.com/doi/10.1002/sim.6472/pdf (http://onlinelibrary.wiley.com/doi/10.1002/sim.6472/pdf).

40. He W, Pinheiro J, Kuznetsova OM. (2014). *Practical Considerations for Adaptive Trial Design and Implementation*. New York, New York: Springer.

41. Bauer P & Köhne K. (1994). Evaluation of experiments with adaptive interim analyses. *Biometrics* 50: 1029–1041, correction in (1996). *Biometrics* 52: 380.

42. Proschan MA & Hunsberger SA. (1995). Designed extension of studies based on conditional power. *Biometrics* 51: 1315–1324.

43. Jennison C & Turnbull BW. (2003). Mid-course sample size modification in clinical trials based on the observed treatment effect. *Statistics in Medicine* 22: 971-993.

44. Mehta CR & Pocock SJ. (2011). Adaptive increase in sample size when interim results are promising: A practical guide with examples. *Statistics in Medicine* 30: 3267–3284.

45. Posch M, Bauer P & Brannath W. (2003). Issues in designing flexible trials. *Statistics in Medicine* 22: 953-969.

46. Lokhnygina Y & Tsiatis AA. (2008). Optimal two-stage group-sequential designs. *Journal of Statistical Planning and Inference* 138: 489-499.

47. Schmitz N. (1993). *Optimal Sequentially Planned Decision Procedures. Lecture Notes in Statistics, Volume 79*. New York, New York: Springer.

48. Wan H, Ellenberg S & Anderson KM. (2015). Stepwise two-stage sample size adaptation. *Statistics in Medicine* 34: 27-38.

49. Mehta CR & Tsiatis AA. (2001). Flexible sample size considerations using information-based interim monitoring. *Drug Information Journal* 35: 1095–1112.

50. Ellenberg SS, Golub H & Mehta C. (2006). Preface to proceedings of workshop adaptive clinical trial designs: ready for prime time? *Statistics in Medicine* 25: 3229–3230.

51. Zelen M. (1969). Play the winner rule and the controlled clinical trial. *Journal of the American Statistical Association* 64: 131-146.

52. Wei LJ & Durham S. (1978). The randomized play-the-winner rule in medical trials. *Journal of the American Statistical Association* 73: 840-843.

53. Hu F & Rosenberger WF. (2006). *The Theory of Response Adaptive Randomization in Clinical Trials*. New York, New York: John Wiley & Sons.

54. Le Tourneau C, Lee JJ & Siu LL. (2009). Dose escalation methods in phase I cancer clinical trials. *Journal of the National Cancer Institute* 101: 708-720.

55. O'Quigley J, Pepe M & Fisher L. (1990). Continual reassessment method: A practical design for phase 1 clinical trials in cancer. *Biometrics* 46: 33-48.

56. Ji Y, Liu P, Li Y & Bekele BN. (2010). A modified toxicity probability interval method for dose-finding trials. *Clinical Trials* 7: 653-663.

57. Maurer W & Bretz F. (2013). Multiple testing in group sequential trials using graphical approaches. *Statistics in Biopharmaceutical Research* 5: 311-320.

58. Sugitani T, Bretz F & Maurer W. (2014). A simple and flexible graphical approach for adaptive group-sequential clinical trials. *Journal of Biopharmaceutical Statistics* DOI: 10.1080/10543406.2014.972509.

59. Antonijevic Z, Bolognese J, Burman CF, Chuang-Stein C, Jennison C, Kimber M, Marchenko O, Patel N & Pinheiro J. (2013). A progress report from the DIA Adaptive Program Work Stream. *Biopharmaceutical Report* 20: 3-9.

60. Gallo P, Chuang-Stein C, Dragalin V, Gaydos B, Krams M & Pinheiro J. (2006). Adaptive designs in clinical drug development – An executive summary of the PhRMA Working Group. *Journal of Biopharmaceutical Statistics* 16: 275-283.

61. Chuang-Stein C, Bretz F, Komiyama O & Quinlan J. (2009). Interactions with regulatory agencies to enhance the understanding and acceptance of adaptive designs. A report by members of the PhRMA Adaptive Design Working Group. *Regulatory Focus* 14: 36-42.

62. Gaydos B, Anderson K, Berry D, Burnham N, Chuang-Stein C, Dudinak J, Fardipour P, Gallo P, Givens S, Lewis R, Maca J, Pinheiro J, Pritchett Y & Krams M. (2009). Good practices for adaptive clinical trials in pharmaceutical product development. *Drug Information Journal* 43: 539-556.

63. Bornkamp B, Bretz F, Dmitrienko A, Enas G, Gaydos B, Hsu CH, Koenig F, Krams M, Liu Q, Neuenschwander B, Parke T, Pinheiro J, Roy A, Sax R & Shen F. (2007). Innovative approaches for designing and analyzing adaptive dose-ranging trials (with discussion). *Journal of Biopharmaceutical Statistics* 17: 965–995.

64. Thomas N, Sweeney K & Somayaji V. (2014). Meta-Analysis of clinical dose–response in a large drug development portfolio. *Statistics in Biopharmaceutical Research* 6: 302-317.

65. Thomas N, Roy D, Somayaji V & Sweeney K. (2014). Meta-analyses of clinical dose response. Presentation at the European Medicines Agency/European Federation of Pharmaceutical Industries and Associations workshop on the importance of dose finding and dose selection for the successful development, licensing and lifecycle management of medicinal products. Available at: http://www.ema.europa.eu/docs/

en_GB/document_library/Presentation/2015/01/WC500179795.pdf (http://www.ema.europa.eu/docs/en_GB/document_library/Presentation/2015/01/WC500179795.pdf).

66. Milligan PA, Brown MJ, Marchant B,Martin SW van der Graaf PH, Benson N, Nucci G, Nichols DJ, Boyd RA, Mandema JW, Krishnaswami S, Zwillich S, Gruben D, Anziano RJ, Stock TC & Lalonde RL. (2013). Model-based drug development: A rational approach to efficiently accelerate drug development. *Clinical Pharmacology & Therapeutics* 93: 502-514.
67. Chang M. (2014). *Adaptive Design Theory and Implementation Using SAS and R, Second Edition*. Boca Raton, Florida: CRC Press.
68. Tymofyeyev Y. (2014). A review of available software and capabilities for adaptive designs. In: Weili H, Pinheiro J & Kuznetsova OM, eds. *Practical Considerations for Adaptive Trial Design and Implementation*. New York, New York: Springer.
69. Antonijevic Z, Gallo P, Chuang-Stein C, Dragalin V, Loewy J, Menon S, Miller E, Morgan CC & Sanchez M. (2013). Views on emerging issues pertaining to data monitoring committees for adaptive trials. *Therapeutic Innovation & Regulatory Science* 47: 495-502.
70. DiMasi J. (2014). Cost to develop a drug more than doubles to $2.56 billion. Available at: http://www.bloomberg.com/news/2014-11-18/cost-to-develop-a-drug-more-than-doubles-to-2-56-billion.html (http://www.bloomberg.com/news/2014-11-18/cost-to-develop-a-drug-more-than-doubles-to-2-56-billion.html).
71. U.S. Food and Drug Administration. (2008). Guidance to Industry: Diabetes Mellitus — Evaluating Cardiovascular Risk in New Antidiabetic Therapies to Treat Type 2 Diabetes. Available at: http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm071627.pdf (http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm071627.pdf).
72. Gregg EW, Zhuo X, Cheng YJ, Albright AL, Venkat Narayan KM & Thompson TJ. (2014). Trends in lifetime risk and years of life lost due to diabetes in the USA, 1985–2011: a modelling study. *The Lancet Diabetes & Endocrinology* 2: 867-874.
73. Cummings JL, Morstorf T, and Zhong K. (2014). Alzheimer's disease drug-development pipeline: few candidates, frequent failures. *Alzheimer's Research & Therapy* 6:37. Available at: http://alzres.com/content/6/4/37 (http://alzres.com/content/6/4/37).
74. Gulfo JV. (2015, January 8). FDA 2014 approvals – the message behind the numbers. The Hill. Available at: http://thehill.com/blogs/congress-blog/healthcare/228803-fda-2014-approvals-the-message-behind-the-numbers (http://thehill.com/blogs/congress-blog/healthcare/228803-fda-2014-approvals-the-message-behind-the-numbers).
75. Barker AD, Sigman CC, Kelloff GJ, Hylton NM, Berry DA & Esserman LJ. (2009). I-SPY 2: An adaptive breast cancer trial design in the setting of neoadjuvant chemotherapy. *Clinical Pharmacology & Therapeutics* 86: 97-100.
76. NCI Press Release. (2014, June 16). Lung-MAP launches: First precision medicine trial from national clinical trials network. Available at: http://www.cancer.gov/newscenter/newsfromnci/2014/LungMAPlaunch (http://www.cancer.gov/newscenter/newsfromnci/2014/LungMAPlaunch).
77. President's Council of Advisors on Science and Technology. (2014). Report to the President on Combating Antibiotic Resistance. Available at: http://www.whitehouse.gov/sites/default/files/microsites/ostp/PCAST/

pcast_carb_report_sept2014.pdf (http://www.whitehouse.gov/sites/default/files/microsites/ostp/PCAST/pcast_carb_report_sept2014.pdf).

78. Murphy SA, Oslin DW, Rush AJ, Zhu J for MCATS. (2007). Methodological challenges in constructing effective treatment sequences for chronic psychiatric disorders. *Neuropsychopharmacology* 32: 257-262.

79. Ivanova A, Qaqish B & Schoenfeld D. (2011). Optimality, sample size and power calculations for the sequential parallel comparison design. *Statistics in Medicine* 30: 2793-2803.

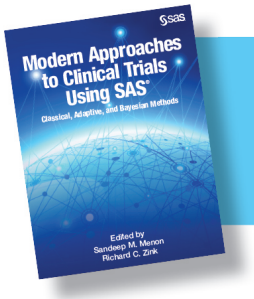
80. Elsäßer A, Regnstrom J, Vetter T, Koenig F, Hemmings RJ, Greco M, Papaluca-Amati M & Posch M. (2014). Adaptive clinical trial designs for European marketing authorization: a survey of scientific advice letters from the European Medicines Agency. *Trials* 15: 383. Available at http://www.trialsjournal.com/content/15/1/383 (http://www.trialsjournal.com/content/15/1/383).

81. Burman CF & Chuang-Stein C. (2009, May 21). Adaptive designs: A fad or the future of clinical research? Applied Clinical Trials. Available at: http://appliedclinicaltrialsonline.findpharma.com/appliedclinicaltrials/article/articleDetail.jsp?id=598938&sk=&date=&pageID=2 (http://appliedclinicaltrialsonline.findpharma.com/appliedclinicaltrials/article/articleDetail.jsp?id=598938&sk=&date=&pageID=2).

Authors

Christy Chuang-Stein, Chuang-Stein Consulting, Kalamazoo, MI, USA.

Keaven M. Anderson, Executive Director, Late Development Statistics, Merck Research Laboratories, North Wales, PA, USA.



Index

B

C

D

E

F

G

H

I

J

K

L

M

N

O

P

R

S

T

U

About This Book

Purpose

Modern Approaches to Clinical Trials Using SAS®: Classical, Adaptive, and Bayesian Methods is unique and multifaceted, covering several domains of modern clinical trial design, including classical, group sequential, adaptive, and Bayesian methods that are applicable to and widely used in various phases of pharmaceutical development. Topics covered include, but are not limited to, dose-response and dose-escalation designs; sequential methods to stop trials early for overwhelming efficacy, safety, or futility; Bayesian designs that incorporate historical data; adaptive sample size re-estimation; adaptive randomization to allocate subjects to more effective treatments; and population enrichment designs. Methods are illustrated using clinical trials from diverse therapeutic areas, including dermatology, endocrinology, infectious disease, neurology, oncology, and rheumatology. Individual chapters are authored by renowned contributors, experts, and key opinion leaders from the pharmaceutical/medical device industry or academia.

Numerous real-world examples and sample SAS code enable users to readily apply novel clinical trial design and analysis methodologies in practice.

Is This Book for You?

This book is intended for biostatisticians, pharmacometricians, clinical developers, and statistical programmers involved in the design, analysis, and interpretation of clinical trials. Further, students in graduate and post-graduate programs in statistics or biostatistics will benefit from the many practical illustrations of statistical concepts.

Prerequisites

Based on the above audience, users will benefit most from this book with some graduate training in statistics or biostatistics, and some experience or exposure to clinical trials. Some experience with simulation may be useful, though this is not required to use this book. Some experience with SAS/STAT procedures, SAS/IML, and the SAS macro language is expected.

About the Examples

Software Used to Develop the Book's Content

The output, figures, and examples presented were generated using the third maintenance release of SAS 9.4 (TS1M3), including SAS/STAT 14.1 and SAS/IML 14.1. However, the code has and is expected to generate the appropriate results using earlier releases of SAS.

Example Code and Data

Code is available for download from http://support.sas.com/publishing/authors (select the name of the author); then, look for the cover thumbnail of this book and select Example Code and Data.

Output and Graphics Used in This Book

Figures were generated using SAS and saved as TIF files. Output was captured from HTML using FullShot 9.5 Professional.

Additional Resources

SAS offers the following books for statisticians engaged in clinical trials.

1. Dmitrienko A, Molenberghs G, Chuang-Stein C & Offen W. (2005). *Analysis of Clinical Trials Using SAS®: A Practical Guide*. Cary, North Carolina: SAS Institute Inc.
2. Dmitrienko A, Chuang-Stein C & D'Agostino R. (2007). *Pharmaceutical Statistics Using SAS®: A Practical Guide*. Cary, North Carolina: SAS Institute Inc.
3. Wicklin R. (2013). *Simulating Data with SAS®*. Cary, North Carolina: SAS Institute Inc.
4. Zink RC. (2014). *Risk-Based Monitoring and Fraud Detection in Clinical Trials Using JMP® and SAS®*. Cary, North Carolina: SAS Institute Inc.

Keep in Touch

We look forward to hearing from you. We invite questions, comments, and concerns. If you want to contact us about a specific book, please include the book title in your correspondence.

To Contact the Author through SAS Press

By e-mail: saspress@sas.com

Via the Web: http://support.sas.com/author_feedback

SAS Books

For a complete list of books available through SAS, visit http://support.sas.com/bookstore.

Phone: 1-800-727-3228

Fax: 1-919-677-8166

E-mail: sasbook@sas.com

SAS Book Report

Receive up-to-date information about all new SAS publications via e-mail by subscribing to the SAS Book Report monthly eNewsletter. Visit http://support.sas.com/sbr.

Publish with SAS

SAS is recruiting authors! Are you interested in writing a book? Visit http://support.sas.com/saspress for more information.

About the Authors

Sandeep Menon, PhD, is currently the Vice President and Head of the Statistical Research and Consulting Center (SRCC) at Pfizer Inc., and he also holds adjunct faculty positions at Boston University and Tufts University School of Medicine. His group, located at different Pfizer sites globally, provides scientific and statistical leadership, and consultation to the global head of biostatistics, various quantitative groups within Pfizer, senior Pfizer management in discovery, clinical development, legal, commercial and marketing. His responsibilities also include providing a strong presence for Pfizer in regulatory and professional circles to influence content of regulatory guidelines and their interpretation in practice. Previously he held positions of responsibility and leadership where he was in charge of all the biostatistics activities for the entire portfolio in his unit, spanning from discovery (target) through proof-of-concept studies for supporting immunology and autoimmune disease, inflammation and remodeling, rare diseases, cardiovascular and metabolism, and center of therapeutic innovation. He was responsible for overseeing statistical aspects of more than 40 clinical trials, over 25 compounds, and 20 indications. He is a core member of the Global Statistics and Triad (Statistics, Clinical and Clinical Pharmacology) Leadership team. He has been in the industry for over a decade and prior to joining Pfizer he worked at Biogen Idec, Aptiv Solutions, and Harvard Clinical Research Institute. He is very passionate about adaptive designs and personalized medicine. He is the coauthor and coeditor of *Clinical and Statistical Considerations in Personalized Medicine* (2014). He is an active member of the American Statistical Association (ASA), serving as a committee member for the prestigious ASA Samuel S. Wilks Memorial Award. He is the co-chair of the DIA-sponsored sub-team on personalized medicine, core member in the DIA working group for small populations, and an invited program committee member at the Biopharmaceutical Applied Statistics Symposium (BASS). He received his medical degree from Bangalore (Karnataka) University, India, and later completed his master's and PhD in Biostatistics at Boston University.

Richard C. Zink, PhD, is Principal Research Statistician Developer in the JMP Life Sciences division at SAS Institute. He is currently a developer for JMP Clinical, an innovative software package designed to streamline the review of clinical trial data. He joined SAS in 2011 after eight years in the pharmaceutical industry, where he designed and analyzed clinical trials in a variety of therapeutic areas, participated in US and European drug submissions, and two FDA advisory committee hearings. He is an active member of the Biopharmaceutical Section of the American Statistical Association (ASA), serving as industry co-chair for the 2015 ASA Biopharmaceutical Section Statistics Workshop, and as a member of the Safety Scientific Working Group. He is a member of the Drug Information Association where he serves as Statistics Section Editor for *Therapeutic Innovation & Regulatory Science*. Richard is a member of Statisticians in the Pharmaceutical Industry, and holds a PhD in Biostatistics from the University of North

Carolina at Chapel Hill, where he serves as an adjunct faculty member. He is author of *Risk-Based Monitoring and Fraud Detection in Clinical Trials Using JMP® and SAS®*.