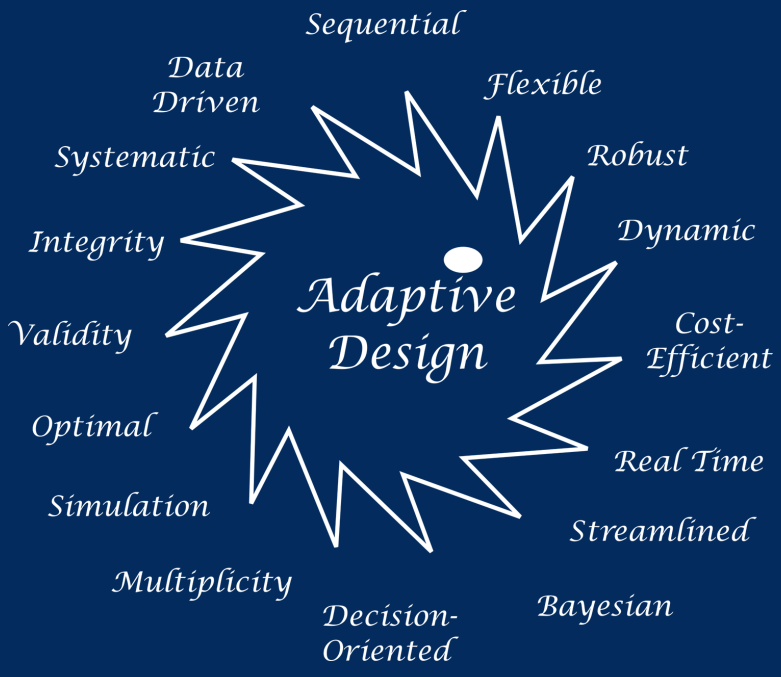
Chapman & Hall/CRC Biostatistics Series

Adaptive Design Theory and Implementation Using SAS and R

Second Edition



Mark Chang

CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK

Adaptive Design Theory and Implementation Using SAS and R

Second Edition

Chapman & Hall/CRC Biostatistics Series

Published Titles

Adaptive Design Methods in Clinical Trials, Second Edition
Shein-Chung Chow and Mark Chang

Adaptive Design Theory and Implementation Using SAS and R, Second Edition
Mark Chang

Advanced Bayesian Methods for Medical Test Accuracy
Lyle D. Broemeling

Advances in Clinical Trial Biostatistics
Nancy L. Geller

Applied Meta-Analysis with R
Ding-Geng (Din) Chen and Karl E. Peace

Basic Statistics and Pharmaceutical Statistical Applications, Second Edition
James E. De Muth

Bayesian Adaptive Methods for Clinical Trials
Scott M. Berry, Bradley P. Carlin, J. Jack Lee, and Peter Muller

Bayesian Analysis Made Simple: An Excel GUI for WinBUGS
Phil Woodward

Bayesian Methods for Measures of Agreement
Lyle D. Broemeling

Bayesian Methods in Epidemiology
Lyle D. Broemeling

Bayesian Methods in Health Economics
Gianluca Baio

Bayesian Missing Data Problems: EM, Data Augmentation and Noniterative Computation
Ming T. Tan, Guo-Liang Tian, and Kai Wang Ng

Bayesian Modeling in Bioinformatics
Dipak K. Dey, Samiran Ghosh, and Bani K. Mallick

Benefit-Risk Assessment in Pharmaceutical Research and Development
Andreas Sashegyi, James Felli, and Rebecca Noel

Biosimilars: Design and Analysis of Follow-on Biologics
Shein-Chung Chow

Biostatistics: A Computing Approach
Stewart J. Anderson

Causal Analysis in Biomedicine and Epidemiology: Based on Minimal Sufficient Causation
Mikel Aickin

Clinical and Statistical Considerations in Personalized Medicine
Claudio Carini, Sandeep Menon, and Mark Chang

Clinical Trial Data Analysis using R
Ding-Geng (Din) Chen and Karl E. Peace

Clinical Trial Methodology
Karl E. Peace and Ding-Geng (Din) Chen

Computational Methods in Biomedical Research
Ravindra Khattree and Dayanand N. Naik

Computational Pharmacokinetics
Anders Källén

Confidence Intervals for Proportions and Related Measures of Effect Size
Robert G. Newcombe

Controversial Statistical Issues in Clinical Trials
Shein-Chung Chow

Data and Safety Monitoring Committees in Clinical Trials
Jay Herson

Design and Analysis of Animal Studies in Pharmaceutical Development
Shein-Chung Chow and Jen-pei Liu

Design and Analysis of Bioavailability and Bioequivalence Studies, Third Edition
Shein-Chung Chow and Jen-pei Liu

Design and Analysis of Bridging Studies
Jen-pei Liu, Shein-Chung Chow, and Chin-Fu Hsiao

Design and Analysis of Clinical Trials with Time-to-Event Endpoints
Karl E. Peace

Design and Analysis of Non-Inferiority Trials
Mark D. Rothmann, Brian L. Wiens, and Ivan S. F. Chan

Difference Equations with Public Health Applications
Lemuel A. Moyé and Asha Seth Kapadia

DNA Methylation Microarrays: Experimental Design and Statistical Analysis
Sun-Chong Wang and Arturas Petronis

DNA Microarrays and Related Genomics Techniques: Design, Analysis, and Interpretation of Experiments
David B. Allison, Grier P. Page, T. Mark Beasley, and Jode W. Edwards

Dose Finding by the Continual Reassessment Method
Ying Kuen Cheung

Elementary Bayesian Biostatistics
Lemuel A. Moyé

Frailty Models in Survival Analysis
Andreas Wienke

Generalized Linear Models: A Bayesian Perspective
Dipak K. Dey, Sujit K. Ghosh, and Bani K. Mallick

Handbook of Regression and Modeling: Applications for the Clinical and Pharmaceutical Industries
Daryl S. Paulson

Inference Principles for Biostaticians
Ian C. Marschner

Interval-Censored Time-to-Event Data: Methods and Applications
Ding-Geng (Din) Chen, Jianguo Sun, and Karl E. Peace

Joint Models for Longitudinal and Time-to-Event Data: With Applications in R
Dimitris Rizopoulos

Measures of Interobserver Agreement and Reliability, Second Edition
Mohamed M. Shoukri

Medical Biostatistics, Third Edition
A. Indrayan

Meta-Analysis in Medicine and Health Policy
Dalene Stangl and Donald A. Berry

Mixed Effects Models for the Population Approach: Models, Tasks, Methods and Tools
Marc Lavielle

Monte Carlo Simulation for the Pharmaceutical Industry: Concepts, Algorithms, and Case Studies
Mark Chang

Multiple Testing Problems in Pharmaceutical Statistics
Alex Dmitrienko, Ajit C. Tamhane, and Frank Bretz

Noninferiority Testing in Clinical Trials: Issues and Challenges
Tie-Hua Ng

Optimal Design for Nonlinear Response Models
Valerii V. Fedorov and Sergei L. Leonov

Patient-Reported Outcomes: Measurement, Implementation and Interpretation
Joseph C. Cappelleri, Kelly H. Zou, Andrew G. Bushmakin, Jose Ma. J. Alvir, Demissie Alemayehu, and Tara Symonds

Quantitative Evaluation of Safety in Drug Development: Design, Analysis and Reporting
Qi Jiang and H. Amy Xia

Randomized Clinical Trials of Nonpharmacological Treatments
Isabelle Boutron, Philippe Ravaud, and David Moher

Randomized Phase II Cancer Clinical Trials
Sin-Ho Jung

Sample Size Calculations for Clustered and Longitudinal Outcomes in Clinical Research
Chul Ahn, Moonseoung Heo, and Song Zhang

Sample Size Calculations in Clinical Research, Second Edition
Shein-Chung Chow, Jun Shao and Hansheng Wang

Statistical Analysis of Human Growth and Development
Yin Bun Cheung

Statistical Design and Analysis of Stability Studies
Shein-Chung Chow

Statistical Evaluation of Diagnostic Performance: Topics in ROC Analysis
Kelly H. Zou, Aiyi Liu, Andriy Bandos, Lucila Ohno-Machado, and Howard Rockette

Statistical Methods for Clinical Trials
Mark X. Norleans

Statistics in Drug Research: Methodologies and Recent Developments
Shein-Chung Chow and Jun Shao

Statistics in the Pharmaceutical Industry, Third Edition
Ralph Buncher and Jia-Yeong Tsay

Survival Analysis in Medicine and Genetics
Jialiang Li and Shuangge Ma

Theory of Drug Development
Eric B. Holmgren

Translational Medicine: Strategies and Statistical Methods
Dennis Cosmatos and Shein-Chung Chow

Chapman & Hall/CRC Biostatistics Series

Adaptive Design Theory and Implementation Using SAS and R

Second Edition

Mark Chang
AMAG Pharmaceuticals, Inc.
Lexington, Massachusetts, USA

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

A CHAPMAN & HALL BOOK

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Version Date: 20141022

International Standard Book Number-13: 978-1-4822-5660-4 (eBook - PDF)

To those who are striving toward a better way

Series Introduction

The primary objectives of the Biostatistics Book Series are to provide useful reference books for researchers and scientists in academia, industry, and government, and also to offer textbooks for undergraduate and/or graduate courses in the area of biostatistics. This book series will provide comprehensive and unified presentations of statistical designs and analyses of important applications in biostatistics, such as those in biopharmaceuticals. A well-balanced summary will be given of current and recently developed statistical methods and interpretations for both statisticians and researchers/scientists with minimal statistical knowledge who are engaged in the field of applied biostatistics. The series is committed to providing easy-to-understand, state-of-the-art references and textbooks. In each volume, statistical concepts and methodologies will be illustrated through real world examples.

In the past several decades, it has been recognized that increasing spending of biomedical research does not reflect an increase of the success rate of pharmaceutical (clinical) development. As a result, the United States Food and Drug Administration (FDA) kicked off a Critical Path Initiative to assist the sponsors in identifying the scientific challenges underlying the medical product pipeline problems. In 2006, the FDA released a Critical Path Opportunities List that outlines 76 initial projects (six broad topic areas) to bridge the gap between the quick pace of new biomedical discoveries and the slower pace at which those discoveries are currently developed into therapies. Among the 76 initial projects, the FDA calls for advancing innovative trial designs, especially for the use of prior experience or accumulated information in trial design. Many researchers interpret it as the encouragement for the use of adaptive design methods in clinical trials.

In clinical trials, it is not uncommon to modify trial and/or statistical procedures during the conduct of the trials based on the review of interim

data. The purpose is not only to efficiently identify clinical benefits of the test treatment under investigation, but also to increase the probability of success of clinical development. The use of adaptive design methods for modifying the trial and/or statistical procedures of on-going clinical trials based on accrued data has been practiced for years in clinical research. However, it is a concern whether the p-value or confidence interval regarding the treatment effect obtained after the modification is reliable or correct. In addition, it is also a concern that the use of adaptive design methods in a clinical trial may lead to a totally different trial that is unable to address scientific/medical questions that the trial is intended to answer. In their book, Chow and Chang (2006) provide a comprehensive summarization of statistical methods for the use of adaptive design methods in clinical trials. This volume provides useful approaches for implementation of adaptive design methods in clinical trials through the application of statistical software such as SAS and R. It covers statistical methods for various adaptive designs such as adaptive group sequential design, adaptive dose-escalation design, adaptive seamless phase-II/III trial design (drop-the-losers design), and biomarker-adaptive design. It would be beneficial to practitioners such as biostatisticians, clinical scientists, and reviewers in regulatory agencies who are engaged in the areas of pharmaceutical research and development.

Shein-Chung Chow
Editor-in-Chief

Preface to the Second Edition

There have been remarkable advancements in methodological study and application of adaptive trials since the publication of the first edition in 2007. I have been thinking about the revision for years and finally I complete the revision today.

In this revision, I have added 12 new chapters, including Chapter 6, Adaptive Noninferiority Design with Paired Binary Data; Chapter 7, Adaptive Design with Incomplete Paired Data; Chapter 12, Blinded and Semi-Blinded Sample-Size Reestimation Design; Chapter 13, Adaptive Design with Coprimary Endpoint; Chapter 15, Pick-the-Winners Design; Chapter 16, The Add-Arm Design for Unimodal Response; Chapter 18, Biomarker-Informed Adaptive Design; Chapter 23, Bayesian Design for Efficacy-Toxicity Trade-Off and Drug Combination; Chapter 24, Bayesian Approach to Biosimilarity Trial; Chapter 25, Adaptive Multiregional Trial Design; Chapter 26, SAS and R Modules for Group Sequential Design; and Chapter 27, Data Analysis of Adaptive Trial.

I have also made major changes to the following chapters: For Chapter 8, K-Stage Adaptive Designs, analytical methods in addition to the simulation methods are now included. For Chapter 11, Unblinded Sample-Size Reestimation Design, the focus is on the comparisons between and discussions on different methods using simulations. I have completely rewritten Chapter 14, Multiple-Endpoint Adaptive Design and Chapter 19, Survival Modeling and Adaptive Treatment Switching, using analytical methods instead of simulation methods. Sequential parallel designs with rerandomization are added in Chapter 20, Response-Adaptive Allocation Design. For Chapter 22, Adaptive Dose-Escalation Trial, I have included the skeleton approach. In the Appendices, some utility SAS code and SAS macros for the add-arm designs are included, and the modified R function for CRM to include the skeleton approach is also provided. In this revision, we have

added nearly 20 new SAS macros and R functions. We have enhanced the exercises or problems in end of each chapter. We want to remind readers that some of the exercises are different from those you would find in a typical textbook of elementary statistics, where all necessary information for solving the problem is exactly given, no more or no less. Some exercises in the book often mimic practical situations, you might be given only the basic information to solve the problem, you need to figure out which information is necessary, what kind of information is missing, and where to get it or how to make assumptions. Those exercises are helpful before you design a real life adaptive trial.

I hope with these revisions and enhancements, readers will find the book useful in designing adaptive trials.

I want to thank Dr. Sandeep Menon for using this book and providing me valuable feedback. I very much appreciate my students, Dr. Jing Wang, Dr. Joseph Wu, Mr. Mike Pickard, Mr. Zhaoyang Teng, and Dr. Yansong Cheng for their creative thinking and hard work. Their contributions are reflected in various chapters. I also thank students in my adaptive design class at Boston University for their engagement and feedback, and thanks to Dr. Sandeep Menon for co-teaching the class with me.

Mark Chang

Preface to the First Edition

This book is about adaptive clinical trial design and computer implementation. Compared to a classical trial design with static features, an adaptive design allows for changing or modifying the characteristics of a trial based on cumulative information. These modifications are often called adaptations. The word *adaptation* is so familiar to us because we constantly make adaptations in our daily lives according to what we learn over time. Some of the adaptations are necessary for survival, while others are made to improve our quality of life. We should be equally smart in conducting clinical trials by making adaptations based on what we learn as the trial progresses. These adaptations are made because they can improve the efficiency of the trial design, provide earlier remedies, and reduce the time and cost of drug development. An adaptive design is also ethically important. It allows for stopping a trial earlier if the risk to subjects outweighs the benefit, or when there is early evidence of efficacy for a safe drug. An adaptive design may allow for randomizing more patients to the superior treatment arms and reducing exposure to inefficacious, but potentially toxic, doses. An adaptive design can also be used to identify better target populations through early biomarker responses.

The aims of this book are to provide a unified and concise presentation of adaptive design theories, furnish the reader with computer programs in SAS and R (also available at www.statisticians.org) for the design and simulation of adaptive trials, and offer (hopefully) a quick way to master the different adaptive designs through examples that are motivated by real issues in clinical trials. The book covers broad ranges of adaptive methods with an emphasis on the relationships among different methods. As Dr. Simon Day pointed out, there are good and bad adaptive designs; a design is not necessarily good just because it is adaptive. There are many rules and issues that must be considered when implementing adaptive designs.

This book has included most current regulatory views as well as discussions of challenges in planning, execution, analysis, and reporting for adaptive designs.

From a "big picture" view, drug development is a sequence of decision processes. To achieve ultimate success, we cannot consider each trial as an isolated piece; instead, a drug's development must be considered an integrated process, using Bayesian decision theory to optimize the design or program as explained in Chapter 21. It is important to point out that every action we take at each stage of drug development is not with the intent of minimizing the number of errors, but minimizing the impact of errors. For this reason, the power of a hypothesis test is not the ultimate criterion for evaluating a design. Instead, many other factors, such as time, safety, and the magnitude of treatment difference, have to be considered in a utility function. From an even bigger-picture view, we are working in a competitive corporate environment, and statistical game theory will provide the ultimate tool for drug development. In the last chapter of the book, I will pursue an extensive discussion of the controversial issues about statistical theories and the fruitful avenues for future research and application of adaptive designs.

Adaptive design creates a new landscape of drug development. The statistical methodology of adaptive design has been greatly advanced by literature in recent years, and there are an increasing number of trials with adaptive features. The PhRMA and BIO adaptive design working groups have made great contributions in promoting innovative approaches to trial design. In preparing the manuscript of this book, I have benefited from discussions with following colleagues: Shein-Chung Chow, Michael Krams, Donald Berry, Jerry Schindler, Michael Chernick, Bruce Turnbull, Barry Turnbull, Sue-Jane Wang (FDA), Vladimir Dragalin, Qing Liu, Simon Day (MHRA), Susan Kenley, Stan Letovsky, Yuan-Yuan Chiu, Jonca Bull, Gorden Lan, Song Yang, Gang Chen, Meiling Lee, Alex Whitmore, Cyrus Mehta, Carl-Fredrik Burman, Richard Simon, George Chi, James Hung (FDA), Aloka Chakravarty (FDA), Marc Walton (FDA), Robert O'Neill (FDA), Paul Gallo, Christopher Jennison, Jun Shao, Keaven Anderson, Martin Posch, Stuart Pocock, Wassmer Gernot, Andy Grieve, Christy Chung, Jeff Maca, Alun Bedding, Robert Hemmings (MHRA), Jose Pinheiro, Jeff Maca, Katherine Sawyer, Sara Radcliffe, Jessica Oldham, Christian Sonesson, Inna Perevozskaya, Anastasia Ivanova, Brenda Gaydos, Frank Bretz, Wenjin Wang, Suman Bhattacharya, and Judith Quinlan.

I would like to thank Hua Liu, PhD; Hugh Xiao, PhD; Andy Boral, MD; Tracy Zhang, MS; MingXiu Hu, PhD; Alun Bedding, PhD; and Jing Xu, PhD for their careful review and many constructive comments. Thanks to Steve Lewitzky, MS; Kate Rinard, MS; Frank Chen, MS; Hongliang Shi, MS; Tracy Zhang, MS; and Rachel Neuwirth, MS for support. I wish to express my gratitude to the following individuals for sharing their clinical, scientific, and regulatory insights about clinical trials: Andy Boral, MD; Iain Web, MD; Irvin Fox, MD; Jim Gilbert, MD; Ian Walters, MD; Bill Trepicchio, PhD; Mike Cooper, MD; Dixie-Lee Esseltine, MD; Jing Marantz, MD; Chris Webster, and Robert Pietrusko, Pharm D.

Thanks to Jane Porter, MS; Nancy Simonian, MD; and Lisa Aldler, BA for their support during the preparation of this book. Special thanks to Lori Engelhardt, MA, ELS, for careful reviews and many editorial comments.

From Taylor and Francis, I would like to thank David Grubbs, Sunil Nair, Jay Margolis, and Amber Donley for providing me the opportunity to work on this book.

Mark Chang
Millennium Pharmaceuticals, Inc.
Cambridge, Massachusetts, USA
www.statisticians.org

Contents

List of Figures

List of Tables

List of Examples

List of SAS Macros and R Functions

Chapter 1

Introduction

1.1 Motivation

Investment in pharmaceutical research and development has more than doubled in the past decade; however, the increase in spending for biomedical research does not reflect an increased success rate of pharmaceutical development. (Figure 1.1). Reasons for this include the following: (1) a diminished margin for improvement escalates the level of difficulty in proving drug benefits; (2) genomics and other new sciences have not yet reached their full potential; (3) mergers and other business arrangements have decreased candidates; (4) easy targets are the focus as chronic diseases are more difficult to study; (5) failure rates have not improved; and (6) rapidly escalating costs and complexity decrease willingness/ability to bring many candidates forward into the clinic (Woodcock, 2004).

There are several critical areas for improvement in drug development. One of the obvious areas for improvement is the design, conduct, and analysis of clinical trials. Improvement of the clinical trials process includes (1) the development and utilization of biomarkers or genomic markers, (2) the establishment of quantitative disease models, and (3) the use of more informative designs such as adaptive and/or Bayesian designs. In practice, the use of clinical trial simulation, the improvement of clinical trial monitoring, and the adoption of new technologies for prediction of clinical outcome will also help in increasing the probability of success in the clinical development of promising candidates. Most importantly, we should not use the evaluation tools and infrastructure of the last century to develop this century's advances. Instead, an innovative approach using adaptive design methods for clinical development must be implemented.

In the next section, we will provide the definition of adaptive design and brief descriptions of commonly used adaptive designs. In Section 1.2.8,

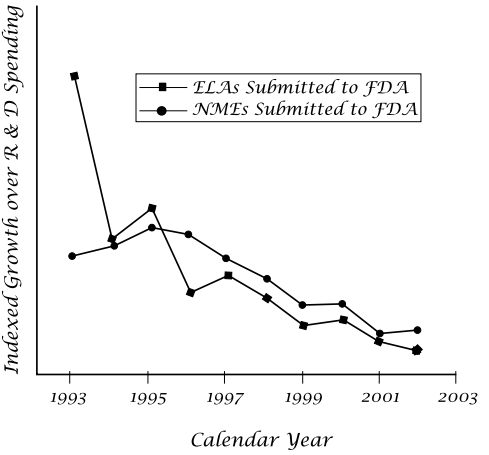


Figure 1.1: Trends in NDAs Submitted to FDA (Data Source: PAREXEXL, 2003)

the importance of computer simulation is discussed. In Section 1.4, we will provide the roadmap for this book.

1.2 Adaptive Design Methods in Clinical Trials

An adaptive design is a clinical trial design that allows adaptations or modifications to aspects of the trial after its initiation without undermining the validity and integrity of the trial (Chang, 2005a; Chow, Chang, and Pong, 2005). The PhRMA Working Group defines an adaptive design as a clinical study design that uses accumulating data to decide how to modify aspects of the study as it continues, without undermining the validity and integrity of the trial (Dragalin, 2006; Gallo et al., 2006).

The adaptations may include, but are not limited to, (1) a group sequential design, (2) an sample-size adjustable design, (3) a drop-losers design, (4) an adaptive treatment allocation design, (5) an adaptive dose-escalation design, (6) a biomarker-adaptive design, (7) an adaptive treatment-switching design, (8) an adaptive dose-finding design, and (9) a combined adaptive design. An adaptive design usually consists of multiple stages. At each stage, data analyses are conducted, and adaptations are taken based on updated information to maximize the probability of success. An adaptive design is also known as a flexible design (EMEA, 2002).

An adaptive design has to preserve the validity and integrity of the trial. The validity includes internal and external validities. *Internal validity* is the degree to which we are successful in eliminating confounding variables and establishing a cause–effect relationship (treatment effect) within the study itself. A study that readily allows its findings to generalize to the population at large has high *external validity*. *Integrity* involves minimizing operational bias, creating a scientifically sound protocol design, adhering firmly to the study protocol and standard operating procedures (SOPs), executing the trial consistently over time and across sites or countries, providing comprehensive analyses of trial data and unbiased interpretations of the results, and maintaining the confidentiality of the data.

1.2.1 *Group Sequential Design*

A group sequential design (GSD) is an adaptive design that allows for premature termination of a trial due to efficacy or futility, based on the results of interim analyses. GSD was originally developed to obtain clinical benefits under economic constraints. For a trial with a positive result, early stopping ensures that a new drug product can be exploited sooner. If a negative result is indicated, early stopping avoids wasting resources. Sequential methods typically lead to savings in sample size, time, and cost when compared with the classical design with a fixed sample size. Interim analyses also enable management to make appropriate decisions regarding the allocation of limited resources for continued development of a promising treatment. GSD is probably one of the most commonly used adaptive designs in clinical trials.

Basically, there are three different types of GSDs: early efficacy stopping design, early futility stopping design, and early efficacy/futility stopping design. If we believe (based on prior knowledge) that the test treatment is very promising, then an early efficacy stopping design should be used. If we are very concerned that the test treatment may not work, an early futility stopping design should be employed. If we are not certain about the magnitude of the effect size, a GSD permitting early stopping for both efficacy and futility should be considered. In practice, if we have a good knowledge regarding the effect size, then a classical design with a fixed sample-size would be more efficient.

1.2.2 *Sample-Size Reestimation Design*

A sample-size reestimation (SSR) design refers to an adaptive design that allows for sample-size adjustment or reestimation based on the review of

interim analysis results (Figure 1.2). The sample-size requirement for a trial is sensitive to the treatment effect and its variability. An inaccurate estimation of the effect size and its variability could lead to an underpowered or overpowered design, neither of which is desirable. If a trial is underpowered, it will not be able to detect a clinically meaningful difference, and consequently could prevent a potentially effective drug from being delivered to patients. On the other hand, if a trial is overpowered, it could lead to unnecessary exposure of many patients to a potentially harmful compound when the drug, in fact, is not effective. In practice, it is often difficult to estimate the effect size and variability because of many uncertainties during protocol development. Thus, it is desirable to have the flexibility to reestimate the sample size in the middle of the trial.

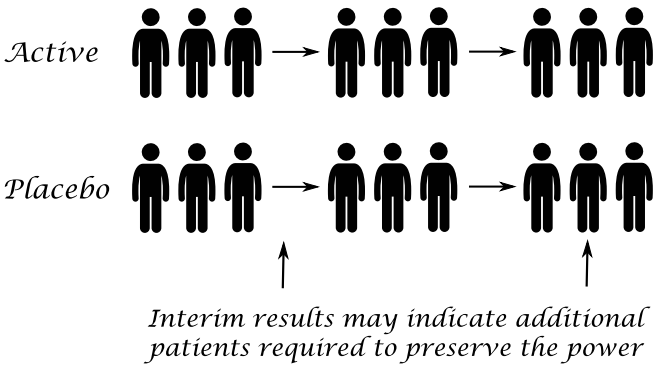


Figure 1.2: Sample-Size Reestimation Design

There are two types of sample-size reestimation procedures, namely, sample-size reestimation based on blinded data and sample-size reestimation based on unblinded data. In the first scenario, the sample adjustment is based on the (observed) pooled variance at the interim analysis to recalculate the required sample size, which does not require unblinding the data. In this scenario, the type-I error adjustment is practically negligible. In the second scenario, the effect size and its variability are reassessed, and sample size is adjusted based on the updated information. The statistical method for adjustment could be based on effect size or the conditional power.

Note that the flexibility in SSR is at the expense of a potential loss of power. Therefore, it is suggested that an SSR be used when there are no good estimates of the effect size and its variability. In the case where there is some knowledge of the effect size and its variability, a classical design would be more efficient.

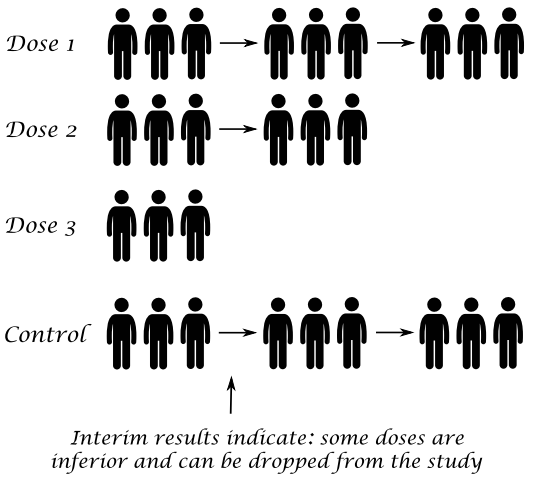


Figure 1.3: Drop-Loser Design

1.2.3 *Drop-Loser Design*

A drop-loser design (DLD) is an adaptive design consisting of multiple stages. At each stage, interim analyses are performed and the losers (i.e., inferior treatment groups) are dropped based on prespecified criteria (Figure 1.3). Ultimately, the best arm(s) are retained. If there is a control group, it is usually retained for the purpose of comparison. This type of design can be used in phase-II/III combined trials. A phase-II clinical trial is often a dose-response study, where the goal is to assess whether there is treatment effect. If there is treatment effect, the goal becomes finding the appropriate dose level (or treatment groups) for the phase-III trials. This type of traditional design is not efficient with respect to time and resources because the phase-II efficacy data are not pooled with data from phase-III trials, which are the pivotal trials for confirming efficacy. Therefore, it is desirable to combine phases II and III so that the data can be used efficiently, and the time required for drug development can be reduced. Bauer and Kieser (1999) provide a two-stage method for this purpose, where investigators can terminate the trial entirely or drop a subset of treatment groups for lack of efficacy after the first stage. As pointed out by Sampson and Sill (2005), the procedure of dropping the losers is highly flexible, and the distributional assumptions are kept to a minimum. However, because of the generality of the method, it is difficult to construct confidence intervals. Sampson and Sill derived a uniformly most powerful, conditionally unbiased test for a normal endpoint.

1.2.4 *Adaptive Randomization Design*

An adaptive randomization/allocation design (ARD) is a design that allows modification of randomization schedules during the conduct of the trial. In clinical trials, randomization is commonly used to ensure a balance with respect to patient characteristics among treatment groups. However, there is another type of ARD, called response-adaptive randomization (RAR), in which the allocation probability is based on the response of the previous patients. RAR was initially proposed because of ethical considerations (i.e., to have a larger probability to allocate patients to a superior treatment group); however, response randomization can be considered a drop-loser design with a seamless allocation probability of shifting from an inferior arm to a superior arm. The well-known response-adaptive models include the randomized play-the-winner (RPW) model (see Figure 1.4), an optimal model that minimizes the number of failures. Other response-adaptive randomizations, such as utility-adaptive randomization, also have been proposed and are combinations of response-adaptive and treatment-adaptive randomization (Chang and Chow, 2005).

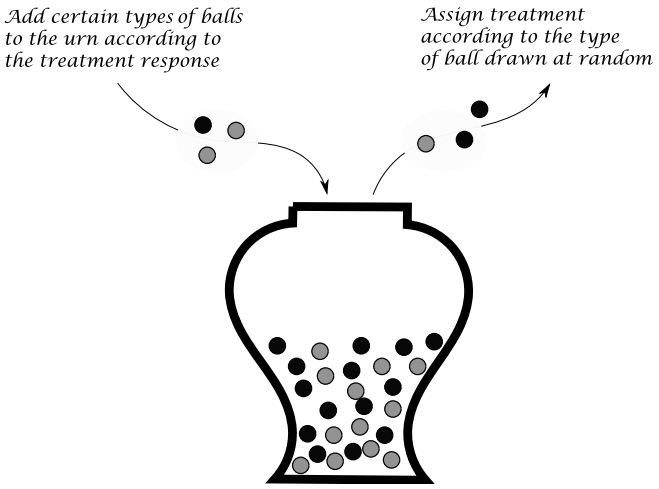


Figure 1.4: Response Adaptive Randomization

1.2.5 *Adaptive Dose-Finding Design*

Dose escalation is often considered in early phases of clinical development for identifying maximum tolerated dose (MTD), which is often considered the optimal dose for later phases of clinical development. An adaptive dose-

finding (or dose-escalation) design is a design in which the dose level used to treat the next-entered patient is dependent on the toxicity of the previous patients, based on some traditional escalation rules (Figure 1.5). Many early dose-escalation rules are adaptive, but the adaptation algorithm is somewhat ad hoc. Recently more advanced dose-escalation rules have been developed using modeling approaches (frequentist or Bayesian framework) such as the continual reassessment method (CRM) (O'Quigley, Pepe, and Fisher, 1990; Chang and Chow, 2005) and other accelerated escalation algorithms. These algorithms can reduce the sample-size and overall toxicity in a trial and improve the accuracy and precision of the estimation of the MTD. Note that CRM can be viewed as a special response-adaptive randomization.

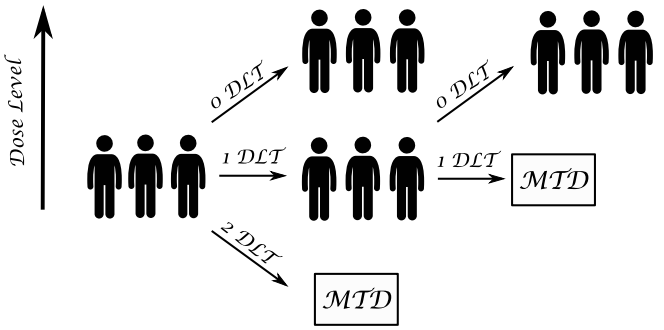


Figure 1.5: Dose Escalation for Maximum Tolerated Dose

1.2.6 *Biomarker-Adaptive Design*

Biomarker-adaptive design (BAD) refers to a design that allows for adaptations using information obtained from biomarkers. A biomarker is a characteristic that is objectively measured and evaluated as an indicator of normal biologic or pathogenic processes or pharmacologic response to a therapeutic intervention (Chakravarty, 2005). A biomarker can be a classifier, prognostic, or predictive marker.

A classifier biomarker is a marker that usually does not change over the course of the study, such as DNA markers. Classifier biomarkers can be used to select the most appropriate target population, or even for personalized treatment. Classifier markers can also be used in other situations. For example, it is often the case that a pharmaceutical company has to make

a decision whether to target a very selective population for whom the test drug likely works well or to target a broader population for whom the test drug is less likely to work well. However, the size of the selective population may be too small to justify the overall benefit to the patient population. In this case, a BAD may be used, where the biomarker response at interim analysis can be used to determine which target populations should be focused on (Figure 1.6).

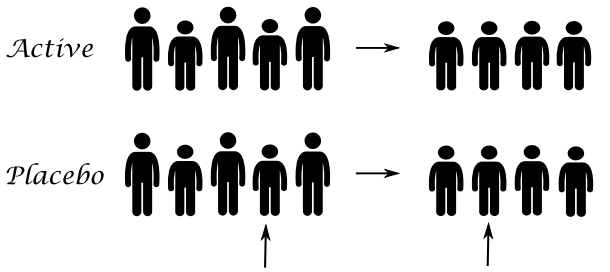


Figure 1.6: Biomarker-Adaptive Design

A prognostic biomarker informs the clinical outcomes, independent of treatment. It provides information about the natural course of the disease in individuals who have or have not received the treatment under study. Prognostic markers can be used to separate good- and poor-prognosis patients at the time of diagnosis. If expression of the marker clearly separates patients with an excellent prognosis from those with a poor prognosis, then the marker can be used to aid the decision about how aggressive the therapy needs to be.

A predictive biomarker informs the treatment effect on the clinical endpoint. Compared to a gold-standard endpoint, such as survival, a biomarker can often be measured earlier, more easily, and more frequently. A biomarker is less subject to competing risks and less affected by other treatment modalities, which may reduce sample size due to a larger effect size. A biomarker could lead to faster decision making. However, validating predictive biomarkers is challenging. BAD simplifies this challenge. In a BAD, "softly" validated biomarkers are used at the interim analysis to assist in decision making, while the final decision can still be based on a gold-standard endpoint, such as survival, to preserve the type-I error (Chang, 2005b).

1.2.7 *Adaptive Treatment-Switching Design*

An adaptive treatment-switching design (ATSD) is a design that allows the investigator to switch a patient's treatment from the initial assignment if there is evidence of lack of efficacy or a safety concern (Figure 1.7).

To evaluate the efficacy and safety of a test treatment for progressive diseases, such as cancers and HIV, a parallel-group, active-control, randomized clinical trial is often conducted. In this type of trial, qualified patients are randomly assigned to receive either an active control (a standard therapy or a treatment currently available in the marketplace) or a test treatment under investigation. Due to ethical considerations, patients are allowed to switch from one treatment to another if there is evidence of lack of efficacy or disease progression. In practice, it is not uncommon that up to 80% of patients may switch from one treatment to another. Sommer and Zeger (1991) referred to the treatment effect among patients who complied with treatment as "biological efficacy." Branson and Whitehead (2002) widened the concept of biological efficacy to encompass the treatment effect as if all patients adhered to their original randomized treatments in clinical studies allowing treatment switching. Despite allowing a switch in treatment, many clinical studies are designed to compare the test treatment with the active control agent as if no patients had ever been switched. This certainly has an impact on the evaluation of the efficacy of the test treatment, because the response-informative switching causes the treatment effect to be confounded. The power for the methods without considering the switching is often lost dramatically because many patients from two groups have eventually taken the same drugs (Shao, Chang, and Chow, 2005). Currently, more approaches have been proposed, which include mixed exponential mode (Chang, 2006a; Chow and Chang, 2006) and a mixture of the Wiener processes (Lee, Chang, and Whitmore, 2008).

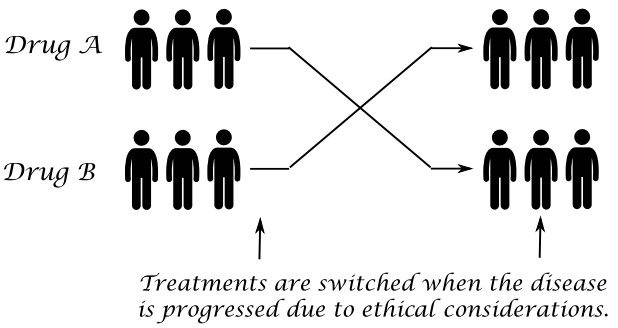


Figure 1.7: Adaptive Treatment Switching

1.2.8 *Clinical Trial Simulation*

Clinical trial simulation (CTS) is a process that mimics clinical trials using computer programs. CTS is particularly important in adaptive designs for several reasons: (1) the statistical theory of adaptive design is complicated with limited analytical solutions available under certain assumptions; (2) the concept of CTS is very intuitive and easy to implement; (3) CTS can be used to model very complicated situations with minimum assumptions, and type-I error can be strongly controlled; (4) using CTS, not only can we calculate the power of an adaptive design, but we can also generate many other important operating characteristics such as expected sample-size, conditional power, and repeated confidence interval—ultimately this leads to the selection of an optimal trial design or clinical development plan; (5) CTS can be used to study the validity and robustness of an adaptive design in different hypothetical clinical settings, or with protocol deviations; (6) CTS can be used to monitor trials, project outcomes, anticipate problems, and suggest remedies before it is too late; (7) CTS can be used to visualize the dynamic trial process from patient recruitment, drug distribution, treatment administration, and pharmacokinetic processes to biomarkers and clinical responses; and finally, (8) CTS has minimal cost associated with it and can be done in a short time.

CTS was started in the early 1970s and became popular in the mid 1990s due to increased computing power. CTS components include (1) a trial Design Mode, which includes design type (parallel, crossover, traditional, adaptive), dosing regimens or algorithms, subject selection criteria, and time, financial, and other constraints; (2) an Execution Model, which models the human behaviors that affect trial execution (e.g., protocol compliance, cooperation culture, decision cycle, regulatory authority, inference of opinion leaders); (3) a Response Model, which includes disease models that imitate the drug behavior (PK and PD models) or intervention mechanism, and an infrastructure model (e.g., timing and validity of the assessment, diagnosis tool); and (4) an Evaluation Model, which includes criteria for evaluating design models, such as utility models and Bayesian decision theory. The CTS model is illustrated in Figure 1.8.

1.2.9 *Regulatory Aspects*

The FDA's Critical Path initiative is a serious attempt to bring attention and focus to the need for targeted scientific efforts to modernize the techniques and methods used to evaluate the safety, efficacy, and quality of

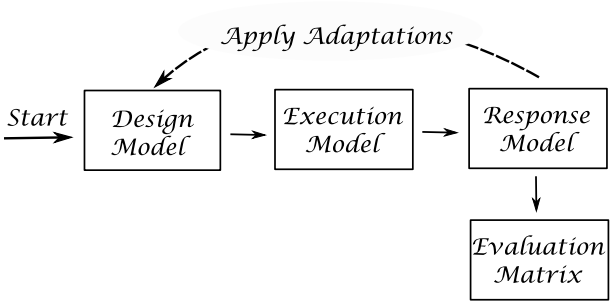


Figure 1.8: Clinical Trial Simulation Model

medical products as they move from product selection and design to mass manufacture. Critical Path is NOT about the drug discovery process. The FDA recognizes that improvement and new technology are needed. The National Institutes of Health (NIH) is getting more involved via the "roadmap" initiative. Critical Path is concerned with the work needed to move a candidate all the way to a marketed product. It is clear that the FDA supports and encourages innovative approaches in drug development. The regulatory agents feel that some adaptive designs are encouraging, but are cautious about others, specially for pivotal studies (EMEA, 2006; Hung, O'Neill, Wang, and Lawrence, 2006; Hung, Wang, and O'Neill, 2006; Temple, 2006).

"Adaptive designs should be encouraged for Phases I and II trials for better exploration of drug effects, whether beneficial or harmful, so that such information can be more optimally used in latter stages of drug development. Controlling false positive conclusions in exploratory phases is also important so that the confirmatory trials in latter stages achieve their goals. The guidance from such trials properly controlling false positives may be more informative to help better design confirmatory trials" (Hung et al., 2006). As pointed out by FDA statistician Dr. Stella Machado, "The two major causes of delayed approval and nonapproval of phase III studies is poor dose selection in early studies and phase III designs [that] don't utilize information from early phase studies" ("The Pink Sheet," Dec. 18, 2006, p. 24). The FDA is granting industry a great deal of leeway in adaptive design in the early learning phase, while at the same time suggesting that emphasis be placed on dose-response and exposure risk. Dr. O'Neill said that learning about the dose-response relationship lies at the heart of adaptive designs. Companies should begin a dialogue about adaptive designs with FDA medical officers and statisticians as early as a year before beginning a trial as suggested by Dr. Robert Powell from the FDA.

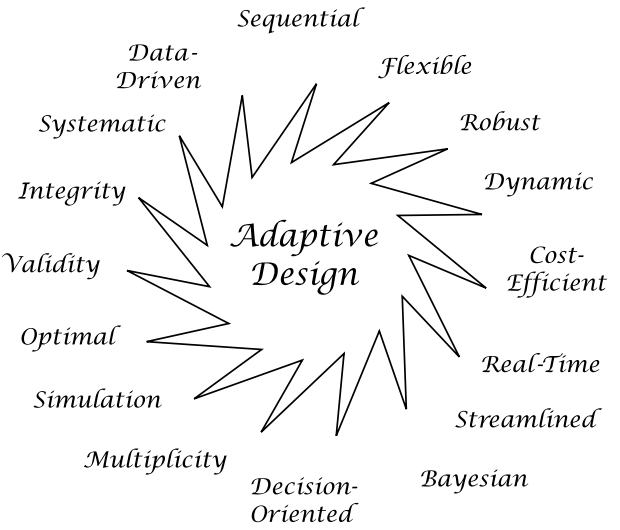


Figure 1.9: Characteristics of Adaptive Designs

1.2.10 *Characteristics of Adaptive Designs*

Adaptive design is a sequential data-driven approach. It is a dynamic process that allows for real-time learning. It is flexible and allows for modifications to the trial, which make the design cost-efficient and robust against the failure. Adaptive design is a systematic way to design different phases of trials, thus streamlining and optimizing the drug development process. In contrast, the traditional approach is composed of weakly connected phase-wise processes. Adaptive design is a decision-oriented, sequential learning process that requires up-front planning and a great deal of collaboration among the different parties involved in the drug development process. To this end, Bayesian methodology and computer simulation play important roles. Finally, the flexibility of adaptive design does not compromise the validity and integrity of the trial or the development process (Figure 1.9).

Adaptive design methods represent a revolution in pharmaceutical research and development. Using adaptive designs, we can increase the chances for success of a trial with a reduced cost. Bayesian approaches provide an ideal tool for optimizing trial designs and development plans. Clinical trial simulations offer a powerful tool to design and monitor trials. Adaptive design, the Bayesian approach, and trial simulation combine to form an ultimate statistical instrument for the most successful drug development programs.

1.3 FAQs about Adaptive Designs

The following questions collected from several journalists from scientific and technological journals (Nature Biotechnology, BioIT World, Contract Pharms, etc.) during the interviews eight years ago are still valuable to discuss today.

1. *What is the classification of an adaptive clinical trial? Is there a consensus in the industry regarding what adaptive trials entail?*

After many conferences and discussions, there is more or less a consensus on the definition of adaptive design. A typical definition is as follows:

An adaptive design is a design that allows modifications to aspects of the trial after its initiation without undermining the validity and integrity of the trial. All adaptive designs involve interim analyses and adaptations or decision making based on the interim results.

There are many ways to classify adaptive designs. The following are common examples of adaptive trials:

- Sample size reestimation design to increase the probability of success
- Early stopping due to efficacy or futility design to reduce cost and time
- Response adaptive randomization design to give patients a better chance of being assigned to superior treatment
- Drop-loser design for adaptive dose finding to reduce sample size by dropping the inferior treatments earlier
- Add-arm design featuring adaptive selection of treatment groups (arms) to reduce the exposure and shorten the study
- Adaptive dose escalation design to minimize toxicity while at the same time acquiring information on maximum tolerated dose
- Adaptive seamless design combining two traditional trials in different phases into a single trial, reducing cost and time to market
- Biomarker enrichment design to have earlier efficacy or safety readout to select better target populations or subpopulation

2. *What challenges does the adaptive trial model present?*

Adaptive designs can reduce time and cost, minimize toxicity, help select the best dose for the patients, and better target populations. With adaptive design, we can develop better science for testing new drugs and, in turn, better science for prescribing them.

There are challenges associated with adaptive design. Statistical methods are available for most common adaptive designs, but for more complicated adaptive designs, the methodologies are still in development.

Operationally, an adaptive design often requires real-time or near real-time data collection and analysis. In this regard, data standardizations, such as CDISC and electronic data capture (EDC), are very helpful in data cleaning and reconciliation. Note that not all adaptive designs require perfectly clean data at interim analysis, but the cleaner the data are, the more efficient the design is. Adaptive designs require the ability to rapidly integrate knowledge and experiences from different disciplines into the decision-making process and, hence, require a shift to a more collaborative working environment among disciplines.

There is no regulatory guidance for adaptive designs at the moment. Adaptive trials are reviewed on a case-by-case basis. Naturally there are fears that a protocol using this innovative approach may be rejected, causing a delay.

The interim unblinding may potentially cause bias and put the integrity of the trial at risk. Therefore, the unblinding procedure should be well established before the trial starts, and frequent unblinding should be avoided. Also, unblinding the premature results to the public could jeopardize the trial.

3. *How would adaptive trials affect traditional phases of drug development? How are safety and efficacy measured in this type of trial?*

Adaptive designs change the way we conduct clinical trials. Trials in different phases can be combined to create a seamless study. The final safety and efficacy requirements are not reduced because of adaptive designs. In fact, with adaptive designs, the efficacy and safety signals are collected and reviewed earlier and more often than in traditional designs. Therefore, we may have a better chance of avoiding unsafe drug exposure to large patient populations. A phase-II and -III combined seamless design, when the trial is carried out to the final stage, has longer-term patient efficacy and safety data than traditional phase-II, phase-III trials; however, precautions should be taken at the interim decision making when data are not mature.

4. *If adaptive trials become widely adopted, how would it impact clinical trial materials and the companies that provide them?*

Depending on the type of adaptive design, there might be requirements for packaging and shipping to be faster and more flexible. Quick and accurate efficacy and safety readouts may also be required. The electronic drug packages with an advanced built-in recording system will be helpful.

If adaptive trials become widely adopted, the drug manufacturers who can provide the materials adaptively will have a better chance of success.

5. *What are some differences between adaptive trials and the traditional trial model with respect to the supply of clinical trial materials?*

For a traditional or classical design, the amount of material required is fixed and can be easily planned before the trial starts. However, for some adaptive trials, the exact amount of required materials is not clear until later stages of the trial. Also the next dosage for a site may not be fully determined until the time of randomization; therefore, vendors may need to develop a better drug distribution strategy.

6. *What areas of clinical development would experience cost/time savings with the adaptive trial model?*

Adaptive design can be used in any phase, even in the preclinical and discovery phases. Drug discovery and development is a sequence of decision processes. The traditional paradigm breaks this into weakly connected fragments or phases. An adaptive approach will eventually be utilized for the whole development process to get the right drug to the right patient at the right time.

Adaptive design may require fewer patients, less trial material, sometimes fewer lab tests, less work for data collection, and fewer data queries to be resolved. However, an adaptive trial requires much more time during up-front planning and simulation studies.

7. *What are some of the regulatory issues that need to be addressed for this type of trial?*

Regulatory documents related to the adaptive clinical trials were issued between 2007 to 2012. They are

(1) European Medicines Agency (EMEA)—Reflection Paper on Methodological Issues in Confirmatory Clinical Trials Planned with an Adaptive Design (October 2007)
(2) U.S. Food and Drug Administration (FDA)—Draft Guidance—Guidance for Industry Adaptive Design Clinical Trials for Drugs and Biologics (February 2010)
(3) U.S. Food and Drug Administration (FDA)—Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials (February 2010)
(4) U.S. Food and Drug Administration (FDA)—Draft Guidance—Guidance for Industry on Enrichment Strategies for Clinical Trials to Support Approval of Human Drugs and Biological Products (December 2012)

If the adaptive design is submitted with solid scientific support and strong ethical considerations and is operationally feasible, there should not

be any fears of rejection of such a design. On the other hand, with a significant increase in adaptive trials in NDA submissions, regulatory bodies may face a temporary shortage of resources for reviewing such designs. Adaptive designs are relatively new to the industry and to regulatory bodies; therefore, there is a lot to learn by doing them. For this reason, it is a good idea to start with adaptive designs in earlier stages of drug development.

1.4 Roadmap

Chapter 2, Classical Design: The classical design and issues raised from the traditional approaches are reviewed. The statistical design methods discussed include one- and two-group designs, multiple-group dose-response designs, as well as equivalence and noninferiority designs.

Chapter 3, Theory of Hypothesis-Based Adaptive Design: Unified theory for adaptive designs, which covers four key statistical elements in adaptive designs: stopping boundary, adjusted p-value, point estimation, and confidence interval is introduced. Discuss how different approaches can be developed under this unified theory and what the common adaptations are.

Chapter 4, Method with Direct Combination of p-values: Using the unified formulation discussed in Chapter 3, the method with an individual stagewise p-value and the methods with the sum and product of the stagewise p-values are discussed in detail for two-stage adaptive designs. Trial examples and step-by-step instructions are provided.

Chapter 5, Method with Inverse-Normal p-values: The inverse-normal method generalizes the classical group sequential method. The method can also be viewed as weighted stagewise statistics and includes several other methods as special cases. Mathematical formulations are derived and examples are provided regarding how to use the method for designing a trial.

Chapter 6, Adaptive Noninferiority Design with Paired Binary Data: Classical and adaptive noninferiority designs with paired binary data are discussed. Examples of sensitivity and specificity studies are provided.

Chapter 7, Adaptive Design with Incomplete Paired Data: When partial paired data is missing, the trial data become a mixture of paired and unpaired data. We discuss how to design an adaptive trial to consider missing paired data.

Chapter 8, K-Stage Adaptive Designs: Chapters 4 and 5 are mainly focused on two-stage adaptive designs because these designs are simple and usually have a closed-form solution. In Chapter 8, we use analytical and

simulation approaches to generalize the methods to K-stage designs using analytical methods, SAS macros, and R functions.

Chapter 9, Conditional Error Function Method and Conditional Power: The conditional error function method is a very general approach. We discuss in particular the Proschan-Hunsberger method and the Muller-Schafer method. We will compare the conditional error functions for various other methods and study the relationships between different adaptive design methods through the conditional error functions and conditional power.

Chapter 10, Recursive Adaptive Design: The recursive two-stage adaptive design not only offers a closed-form solution for K-stage designs, but also allows for very broad adaptations. We first introduce two powerful principles, the error-spending principle and the conditional error principle, from which we further derive the recursive approach. Examples are provided to illustrate the different applications of this method.

Chapter 11, Unblinded Sample-Size Reestimation Design: This chapter is devoted to the commonly used adaptation, unblinded sample-size reestimation. Various sample-size reestimation methods are evaluated and compared. The goal is to demonstrate a way to evaluate different methods under different conditions and to optimize the trial design that fits a particular situation. Practical issues and concerns are also addressed.

Chapter 12, Blinded and Semi-Blinded Sample-Size Reestimation Design: In contrast to unblinded analysis, in this chapter we will discuss the sample-size reestimation without unblinding the treatment code. We will first discuss different methods to estimate the treatment effect without unblinding the randomization code, then discuss the different sample-size reestimation methods. Finally we will see an effective sample size reestimation method with a mixture of blinded and unblinded methods.

Chapter 13, Adaptive Design with Coprimary Endpoint: We will discuss how to control type-I error in an adaptive trial with coprimary endpoints, the stopping boundary, the power, and the conditional power, from both analytically and simulation perspective. R-functions are provided.

Chapter 14, Multiple-Endpoint Adaptive Design: One of the most challenging issues is the multiple-endpoint analysis with adaptive design. We will briefly review the multiplicity issues and commonly used methods in classical trials. Then motivated by an actual adaptive design in an oncology trial, we will discuss the methods for the multiple-endpoint issues with coprimary endpoints in adaptive trials.

Chapter 15, Pick-the-Winners Design: We will first discuss the opportunities for phase-II and -III trials combinations. Two adaptive design

methods will be discussed, the common pick-the-winner design and the adaptive Dunnett test.

Chapter 16, The Add-Arm Design for Unimodal Response: In a classical drop-loser (or drop-arm) design, patients are randomized into all arms (doses) and at the interim analysis, inferior arms are dropped. Therefore, compared to the traditional dose-finding design, this adaptive design can reduce the sample size by not carrying over all doses to the end of the trial or by dropping the losers earlier. However, given a unimodal response, we discuss a more efficient design, the add-arm design.

Chapter 17, Biomarker-Enrichment Design: In this chapter, adaptive design methods are developed for classifier, diagnosis, and predictive markers. SAS macros have been developed for biomarker-adaptive designs. The improvement in efficiency is assessed for difference methods in different scenarios.

Chapter 18, Biomarker-Informed Adaptive Design: The conventional approach uses the patient-level correlation model, together with historical knowledge, to describe the relationship between the biomarker and the primary endpoint. However, this approach ignores the important factor in the relationship between the mean of biomarker response and the primary endpoint; without this consideration, the models turn out to have little effect of biomarker on the primary endpoint. In this chapter, we will discuss a more advanced method that will incorporate the relationships at patient level and the aggregate level.

Chapter 19, Survival Modeling and Adaptive Treatment Switching: Response-adaptive treatment switching and crossover are statistically challenging. Treatment switching is not required for the statistical efficacy of a trial design; rather, it is motivated by an ethical consideration. Several methods are discussed, including the time-dependent exponential, the mixed exponential, and a mixture of Wiener models.

Chapter 20, Response-Adaptive Allocation Design: Response-adaptive randomizations/allocations have many different applications. They can be used to reduce the overall sample-size and the number of patients exposed to ineffective or even toxic regimens. We will discuss some commonly used adaptive randomizations, such as randomized play-the-winner. The sequential parallel design with rerandomization is also discussed.

Chapter 21, Introductory Bayesian Approach in Clinical Trial: The philosophical differences between the Bayesian and frequentist approaches are discussed. Through many examples, the two approaches are compared in terms of design, monitoring, analysis, and interpretation of results. More

importantly, how to use Bayesian decision theory to further improve the efficiency of adaptive designs is discussed with examples.

Chapter 22, Adaptive Dose-Escalation Trial: The adaptive dose-finding designs, or dose-escalation designs, are discussed in this chapter. The goals are to reduce the overall sample size and the number of patients exposed to ineffective or even toxic regimens and to increase the precision and accuracy of MTD (maximum tolerated dose) assessment. We will discuss oncology dose-escalation trials with traditional and Bayesian continual reassessment methods

Chapter 23, Bayesian Design for Efficacy-Toxicity Trade-off and Drug Combination: In this chapter, we will study the more complex Bayesian dose-finding models in two dimensions. Either the outcome has two dimensions, efficacy and toxicity, or the treatment has two dimensions, drug combinations.

Chapter 24, Bayesian Approach to Biosimilarity Trial: Unlike small molecule drug products, for which we can make generic versions that contain the exact same active ingredient as the brand-name drug, biological drugs, such as protein, are large molecule products that are generally produced using a living system or organism, and may be manufactured through biotechnology, derived from natural sources, or produced synthetically. Following the FDA's stepwise totality evidence approach, we will discuss statistical methods and designs that combine different sources of information to provide the totality of the evidence for biosimilar drug approval.

Chapter 25, Adaptive Multiregional Trial Design: A global multiregional clinical trial (MRCT) is an international clinical trial conducted in multiple countries with a uniform study protocol. Its goal is to get the drug approval in multiple countries. We will discuss some regulatory requirements, optimal adaptive MRCT design, and the Bayesian approach.

Chapter 26, SAS and R Modules for Group Sequential Design: We introduce the SAS procedures for group sequential designs and discuss simple examples.

Chapter 27, Data Analysis of Adaptive Trial: Data analyses of an adaptive trial include point and confidence parameter estimates, and adjusted p-values. We discuss the controversial issues surrounding these topics and different types of biases and their adjustments.

Chapter 28, Planning, Execution, Analysis, and Reporting: In this chapter, we discuss the logistic issues with adaptive designs. The topics cover planning, monitoring, analysis, and reporting for adaptive trials. It also includes the most concurrent regulatory views and recommendations.

Chapter 29, Debates in Adaptive Designs: We will present very broad discussions of the challenges and controversies presented by adaptive designs from philosophical and statistical perspectives.

Appendix A: Random Number Generation

Appendix B: A Useful Utility

Appendix C: SAS Macros for Add-Arm Designs

Appendix D: Implementing Adaptive Designs in R

Computer Programs

Most adaptive design methods have been implemented and tested in SAS version 9, and major methods have also been implemented in R. These computer programs are compact (often fewer than 50 lines of SAS code) and ready to use. For convenience, electronic versions of the programs have been made available at **www.statisticians.org**.

The SAS code is enclosed in >>**SAS Macro x.x**>> and <<**SAS**<< or in >>**SAS**>> and <<**SAS**<<. R programs are presented in Appendix B.

Problems

1.1 What are the main differences between classical clinical trial design and adaptive trial design?

1.2 Describe the objectives of different adaptive designs and when different types of adaptive designs should be used.

1.3 What challenges may we face when we adopt the adaptive design? Provide some examples for which a classical instead of an adaptive design should be used.

Chapter 2

Classical Design

2.1 Overview of Drug Development

Pharmaceutical medicine uses all the scientific, clinical, statistical, regulatory, and business knowledge available to provide a challenging and rewarding career. On average, it costs about $1.8 billion to take a new compound to market and only one in 10,000 compounds ever reaches the market. There are three major phases of drug development: (1) preclinical research and development, (2) clinical research and development, and (3) after the compound is on the market, a possible "post-marketing" phase.

The preclinical phase represents bench work (in vitro) followed by animal testing, including kinetics, toxicity, and carcinogenicity. An investigational new drug application (IND) is submitted to the FDA seeking permission to begin the heavily regulated process of clinical testing in human subjects. The clinical research and development phase, representing the time from the beginning of human trials to the new drug application (NDA) submission that seeks permission to market the drug, is by far the longest portion of the drug development cycle and can last from 2 to 10 years (Tonkens, 2005).

Clinical trials are usually divided into three phases. The primary objectives of phase I are (1) to determine the metabolism and pharmacological activities of the drug, the side effects associated with increasing dose, and early evidence of effectiveness, and (2) to obtain sufficient information regarding the drug's pharmacokinetics and pharmacological effects to permit the design of well-controlled and scientifically valid phase-II clinical studies (21 CFR 312.21). Unless it is an oncology study, where the maximum tolerated dose (MTD) is primarily determined by a phase-I dose-escalation study, the dose-response or dose-finding study is usually conducted in phase II, and efficacy is usually the main focus. The choice of study design and

study population in a dose-response trial will depend on the phase of development, therapeutic indication under investigation, and severity of the disease in the patient population of interest (ICH Guideline E4, 1994). Phase-III trials are considered confirmative trials.

The FDA does not actually approve the drug itself for sale. It approves the labeling, the package insert. United States law requires truth in labeling, and the FDA ensures that claims that a drug is safe and effective for treatment of a specified disease or condition have, in fact, been proven. All prescription drugs must have labels, and without proof of the truth of its label, a drug may not be sold in the United States.

In addition to mandated conditional regulatory approval and postmarketing surveillance trials, other reasons sponsors may conduct postmarketing trials include comparing their drug with that of competitors, widening the patient population, changing the formulation or dose regimen, or applying a label extension. A simplified view of the NDA is shown in Figure 2.1 (Tonkens, 2005).

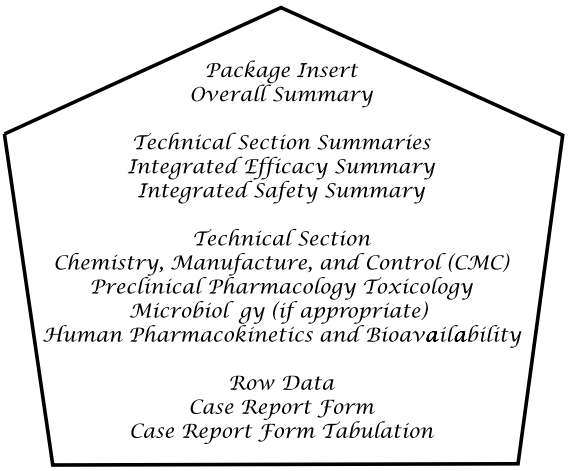


Figure 2.1: A Simplified View of the NDA

In classical trial designs, power and sample-size calculations are a major task. The sample-size calculations for two-group designs have been studied by many scholars, among them Julious (2004), Chow, Shao, and Wang (2003), Machin, et al. (1997), Campbell, Julious, and Altman (1995), and Lachin and Foukes (1986).

In what follows, we will review a unified formulation for sample-size calculation in classical two-arm designs including superiority, noninferiority,

and equivalence trials. We will also discuss some important concepts and issues with the designs that are often misunderstood. We will first discuss two-group superiority and noninferiority designs in Section 2.2. Equivalence studies will be discussed in Section 2.3. Three different types of equivalence studies (average, population, and individual equivalences) are reviewed. We will discuss dose-response studies in Section 2.4. The sample-size calculations for various endpoints are provided based on the contrast test.

2.2 Two-Group Superiority and Noninferiority Designs

2.2.1 *General Approach to Power Calculation*

When testing a null hypothesis $H_0 : \varepsilon \leq 0$ against an alternative hypothesis $H_a : \varepsilon > 0$, where ε is the treatment effect (difference in response), the type-I error rate function is defined as

$$\alpha(\varepsilon) = \Pr\{\text{reject } H_0 \text{ when } H_0 \text{ is true}\}.$$

Note: alternatively, the type-I error rate can be defined as $\sup_{\varepsilon \in H_0} \{\alpha(\varepsilon)\}$. Similarly, the type-II error rate function β is defined as

$$\beta(\varepsilon) = \Pr\{\text{fail to reject } H_0 \text{ when } H_a \text{ is true}\}.$$

For hypothesis testing, knowledge of the distribution of the test statistic under H_0 is required. For sample-size calculation, knowledge of the distribution of the test statistic under a particular H_a is also required. To control the overall type-I error rate at level a constant level α^* under any point of the H_0 domain, the condition $\alpha(\varepsilon) \leq \alpha^*$ for all $\varepsilon \leq 0$ must be satisfied, where α^* is a threshold that is usually larger than 0.025 unless it is a phase-III trial. If $\alpha(\varepsilon)$ is a monotonic function of ε, then the maximum type-I error rate occurs when $\varepsilon = 0$, and the rejection region should be derived under this condition (for this reason we will simply use constant α instead of α^*). For example, for the null hypothesis $H_0 : \mu_2 - \mu_1 \leq 0$, where μ_1 and μ_2 are the means of the two treatment groups, the maximum type-I error rate occurs on the boundary of H_0 when $\mu_2 - \mu_1 = 0$. Let $T = \frac{\hat{\mu}_2 - \hat{\mu}_1}{\hat{\sigma}}$, where $\hat{\mu}_i$ and $\hat{\sigma}$ are the sample mean and pooled sample standard deviation, respectively. Further, let $\Phi_o(T)$ denote the cumulative distribution function (cdf) of the test statistic on the boundary of the null hypothesis domain, and let $\Phi_a(T)$ denote the cdf under H_a. Given this information, under the large sample assumption, $\Phi_o(T)$ is the cdf of the