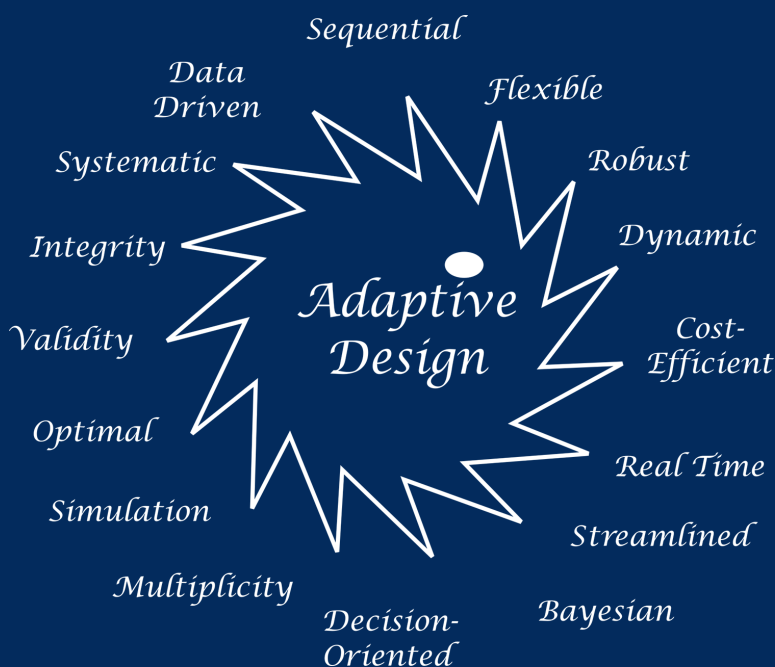


Chapman & Hall/CRC Biostatistics Series

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

Editor-in-Chief

Shein-Chung Chow, Ph.D., Professor, Department of Biostatistics and Bioinformatics,
Duke University School of Medicine, Durham, North Carolina

Series Editors

Byron Jones, Biometrical Fellow, Statistical Methodology, Integrated Information Sciences,
Novartis Pharma AG, Basel, Switzerland

Jen-pei Liu, Professor, Division of Biometry, Department of Agronomy,
National Taiwan University, Taipei, Taiwan

Karl E. Peace, Georgia Cancer Coalition, Distinguished Cancer Scholar, Senior Research Scientist
and Professor of Biostatistics, Jiann-Ping Hsu College of Public Health,
Georgia Southern University, Statesboro, Georgia

Bruce W. Turnbull, Professor, School of Operations Research and Industrial Engineering,
Cornell University, Ithaca, New York

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 Biostatisticians

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. Bushmakina, 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

© 2014 by Taylor & Francis Group, LLC
CRC Press is an imprint of Taylor & Francis Group, an Informa business

No claim to original U.S. Government works
Version Date: 20141022

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

This book contains information obtained from authentic and highly regarded sources. Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access www.copyright.com (<http://www.copyright.com/>) or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Visit the Taylor & Francis Web site at
<http://www.taylorandfrancis.com>

and the CRC Press Web site at
<http://www.crcpress.com>

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	xxix
List of Tables	xxxi
List of Examples	xxxv
List of SAS Macros and R Functions	xxxix
1 Introduction	1
1.1 Motivation	1
1.2 Adaptive Design Methods in Clinical Trials	2
1.2.1 Group Sequential Design	3
1.2.2 Sample-Size Reestimation Design	3
1.2.3 Drop-Loser Design	5
1.2.4 Adaptive Randomization Design	6
1.2.5 Adaptive Dose-Finding Design	6
1.2.6 Biomarker-Adaptive Design	7
1.2.7 Adaptive Treatment-Switching Design	9
1.2.8 Clinical Trial Simulation	10
1.2.9 Regulatory Aspects	10
1.2.10 Characteristics of Adaptive Designs	12
1.3 FAQs about Adaptive Designs	13
1.4 Roadmap	16
2 Classical Design	23
2.1 Overview of Drug Development	23

2.2	Two-Group Superiority and Noninferiority Designs	25
2.2.1	General Approach to Power Calculation	25
2.2.2	Powering Trials Appropriately	29
2.3	Two-Group Equivalence Trial	31
2.3.1	Equivalence Test	31
2.3.2	Average Bioequivalence	34
2.3.3	Population and Individual Bioequivalence	37
2.4	Dose-Response Trials	38
2.4.1	Unified Formulation for Sample-Size	39
2.4.2	Application Examples	41
2.4.3	Determination of Contrast Coefficients	44
2.4.4	SAS Macro for Power and Sample-Size	45
2.5	Summary and Discussion	47
3	Theory of Hypothesis-Based Adaptive Design	53
3.1	Introduction	53
3.2	General Theory	55
3.2.1	Stopping Boundary	55
3.2.2	Formula for Power and Adjusted p -value	56
3.2.3	Selection of Test Statistics	57
3.2.4	Polymorphism	58
3.2.5	Adjusted Point Estimates	59
3.2.6	Derivation of Confidence Intervals	63
3.3	Design Evaluation—Operating Characteristics	64
3.3.1	Stopping Probabilities	64
3.3.2	Expected Duration of an Adaptive Trial	64
3.3.3	Expected Sample Sizes	65
3.3.4	Conditional Power and Futility Index	65
3.3.5	Utility and Decision Theory	66
3.4	Summary	67
4	Method with Direct Combination of p -values	71
4.1	Method Based on Individual p -values	71
4.2	Method Based on the Sum of p -values	76
4.3	Method with Product of p -values	83
4.4	Event-Based Adaptive Design	94
4.5	Adaptive Design for Equivalence Trial	96
4.6	Summary	100
5	Method with Inverse-Normal p -values	103

5.1	Method with Linear Combination of z-Scores	103
5.2	Lehmacher and Wassmer Method	106
5.3	Classical Group Sequential Method	111
5.4	Cui–Hung–Wang Method	114
5.5	Lan–DeMets Method	114
5.5.1	Brownian Motion	115
5.5.2	Lan–DeMets Error-Spending Method	117
5.6	Fisher–Shen Method	120
5.7	Summary	120
6	Adaptive Noninferiority Design with Paired Binary Data	123
6.1	Noninferiority Design	123
6.1.1	Fixed-Margin Method	125
6.1.2	λ -Portion Method	125
6.1.3	Synthesis Method	126
6.2	Noninferiority Design with Fixed-Margin Method for Paired Data	126
6.2.1	Classical Design	126
6.2.2	Adaptive Design	128
6.3	Conditional Power and Sample-Size Reestimation	129
6.4	Type-I Error Control	129
6.5	Prostate Cancer Diagnostic Trial	130
6.5.1	Preliminary Data for Trial Design	130
6.5.2	The Effectiveness Requirements	131
6.5.3	Design for Sensitivity	131
6.5.4	Design for Specificity	132
6.5.5	Summary of Design	134
6.6	Summary	138
7	Adaptive Design with Incomplete Paired Data	141
7.1	Introduction	141
7.2	Mixture of Paired and Unpaired Data	142
7.3	Hypothesis Test	146
7.4	Type-I Error Control and Coverage Probability	146
7.5	Classical Trial Design	147
7.5.1	Power and Sample Size	147
7.5.2	Trial Design Examples	149
7.6	Adaptive Trial Design	151
7.7	Summary	155

8	<i>K</i> -Stage Adaptive Designs	157
8.1	Introduction	157
8.2	Determination of Stopping Boundary	157
8.2.1	Analytical Formulation for MSP	157
8.2.2	Analytical Formulation for MPP	160
8.2.3	Numerical Algorithm For MINP	161
8.3	Monte Carlo Approach	164
8.3.1	Normal Endpoint	164
8.3.2	Binary Endpoint	170
8.3.3	Survival Endpoint	174
8.4	Summary	179
9	Conditional Error Function Method and Conditional Power	183
9.1	Proschan–Hunsberger Method	183
9.2	Denne Method	186
9.3	Müller–Schäfer Method	187
9.4	Extension of Conditional Error Approach	187
9.5	Conditional Power	191
9.6	Adaptive Futility Design	196
9.6.1	Utilization of an Early Futility Boundary	196
9.6.2	Design with a Futility Index	196
9.7	Summary	197
10	Recursive Adaptive Design	199
10.1	p-clud Distribution	199
10.2	Error-Spending and Conditional Error Principles	201
10.3	Recursive Two-Stage Design	203
10.3.1	Sum of Stagewise p -values	204
10.3.2	Product of Stagewise p -values	205
10.3.3	Inverse-Normal Stagewise p -values	206
10.3.4	Application Example	206
10.4	Recursive Combination Tests	209
10.5	Decision Function Method	212
10.6	Summary and Discussion	213
11	Unblinded Sample-Size Reestimation Design	217
11.1	Opportunity	217
11.2	Adaptation Rules	218
11.2.1	Adjustment Based on Effect-Size Ratio	218

11.2.2	Adjustment Based on Conditional Power	219
11.3	Stopping Boundaries for SSR	220
11.4	Basic Considerations in Designing SSR Trial	220
11.5	SAS Macros for Sample-Size Reestimation	225
11.6	Power Comparisons in Promising-Zone	228
11.7	Analysis of Design with Sample-Size Adjustment	233
11.7.1	Adjusted p -value	233
11.7.2	Confidence Interval	234
11.7.3	Adjusted Point Estimates	235
11.8	Trial Examples	236
11.9	Summary and Discussion	240
12	Blinded and Semi-Blinded Sample-Size Reestimation Design	243
12.1	Introduction	243
12.2	Maximum Information Design	244
12.3	Distribution-Informed Design	245
12.3.1	Mean Difference Estimated from Blinded Data	245
12.3.2	Mean Difference: Unequal Variance	247
12.3.3	Rate Difference	249
12.3.4	Semi-Blinded Method	249
12.3.5	Simulation Comparison	250
12.4	Operating Characteristics of Sample-Size Reestimation	253
12.5	Mixed-Method for Sample-Size Reestimation Design	258
12.6	Revealing Treatment Difference via Stratified Randomization	260
12.7	Summary	265
13	Adaptive Design with Coprimary Endpoints	267
13.1	Introduction	267
13.2	Group Sequential Design with Multiple Coprimary Endpoints	269
13.3	Stopping Boundaries	270
13.4	Examples of Coprimary Endpoint Design	272
13.5	Simulation of Group Sequential Trial with Coprimary Endpoints	274
13.6	Conditional Power and Sample Size Reestimation	275
13.7	Summary	280
14	Multiple-Endpoint Adaptive Design	283

14.1	Multiple-Testing Taxonomy	283
14.2	Multiple-Testing Approaches	288
14.2.1	Single-Step Procedures	288
14.2.2	Stepwise Procedures	290
14.2.3	Common Gatekeeper Procedure	295
14.2.4	Tree Gatekeeping Procedure	298
14.3	Multiple-Endpoint Adaptive Design	299
14.3.1	Tang–Geller Method	299
14.3.2	Single Primary with Secondary Endpoints	300
14.4	Summary and Discussion	305
15	Pick-the-Winners Design	307
15.1	Opportunity	307
15.1.1	Impact Overall Alpha Level and Power	307
15.1.2	Reduction in Expected Trial Duration	308
15.1.3	Overview of Multiple-Arm Designs	309
15.2	Pick-the-Winner Design	311
15.3	Adaptive Dunnett Test	315
15.4	Summary and Discussion	317
16	The Add-Arm Design for Unimodal Response	319
16.1	Introduction	319
16.2	The Add-Arm Design	323
16.2.1	Design Description	323
16.2.2	The Interim and Final Test Statistics	325
16.3	Extension of Add-Arm Designs	333
16.3.1	The 5+1 Add-Arm Design	333
16.3.2	The 6+1 Add-Arm Design	335
16.3.3	The 7+1 Add-Arm Design	336
16.4	Comparison of Adaptive Design Methods	338
16.4.1	Threshold c_R and Stopping Boundary c_α	338
16.4.2	Comparison of Seamless Designs	339
16.4.3	Comparisons of Phase-II Dose-Finding Designs	342
16.5	Clinical Trial Examples	343
16.6	Summary	348
17	Biomarker-Enrichment Design	351
17.1	Introduction	351
17.2	Design with Classifier Biomarker	353

17.2.1	Setting the Scene	353
17.2.2	Classical Design with Classifier Biomarker	355
17.2.3	Adaptive Design with Classifier Biomarker	358
17.3	Challenges in Biomarker Validation	362
17.3.1	Classical Design with Biomarker Primary Endpoint	362
17.3.2	Treatment-Biomarker-Endpoint Relationship	362
17.3.3	Multiplicity and False Positive Rate	364
17.3.4	Validation of Biomarkers	365
17.3.5	Biomarkers in Reality	366
17.4	Adaptive Design with Prognostic Biomarkers	366
17.4.1	Optimal Design	366
17.4.2	Prognostic Biomarker in Designing Survival Trial	367
17.5	Adaptive Design with Predictive Marker	368
17.6	Summary and Discussion	368
18	Biomarker-Informed Adaptive Design	371
18.1	Introduction	371
18.2	Motivations and Concepts	372
18.3	Issues in Conventional One-Level Correlation Model	373
18.4	Two-Stage Winner Design	375
18.5	Two-Level Relationship Model for Two-Stage Winner Design	376
18.5.1	Test Statistic and Its Distribution	376
18.5.2	Type-I Error Rate Control	378
18.5.3	Performance Evaluation of Two-Stage Winner Design	379
18.5.4	Parameter Estimation	381
18.6	Hierarchical Model	382
18.7	Summary	385
19	Survival Modeling and Adaptive Treatment Switching	389
19.1	Introduction to Survival Data Modeling	389
19.1.1	Basic Terms in Survival Analysis	389
19.1.2	Maximum Likelihood Method	390
19.1.3	Overview of Survival Model	391
19.2	First-Hitting-Time Model	395
19.2.1	Wiener Process and First Hitting Time	395
19.2.2	Covariates and Link Function	396

19.2.3	Parameter Estimation and Inference	396
19.2.4	Applications of First-Hitting-Time Model	397
19.2.5	Multivariate Model with Biomarkers	398
19.3	Multistage Model	401
19.3.1	General Framework of Multistage Model	401
19.3.2	Covariates and Treatment Switching	403
19.4	Summary and Discussion	405
20	Response-Adaptive Allocation Design	409
20.1	Opportunities	409
20.2	Traditional Randomization Methods	410
20.3	Basic Response-Adaptive Randomizations	411
20.3.1	Play-the-Winner Model	411
20.3.2	Randomized Play-the-Winner Model	412
20.3.3	Optimal Randomized Play-the-Winner	413
20.4	Adaptive Design with Randomized Play-the-Winner	414
20.5	General Response-Adaptive Randomization	418
20.5.1	SAS Macro for K -Arm RAR with Binary Endpoint	418
20.5.2	SAS Macro for K -Arm RAR with Normal Endpoint	420
20.5.3	RAR for General Adaptive Designs	423
20.6	Sequential Parallel Comparison Design	423
20.7	Summary and Discussion	425
21	Introductory Bayesian Approach in Clinical Trial	427
21.1	Introduction	427
21.2	Bayesian Learning Mechanism	428
21.3	Bayesian Basics	429
21.3.1	Bayes' Rule	429
21.3.2	Conjugate Family of Distributions	431
21.4	Trial Design	432
21.4.1	Bayesian for Classical Design	432
21.4.2	Bayesian Power	434
21.4.3	Frequentist Optimization	435
21.4.4	Bayesian Optimal Adaptive Designs	437
21.5	Trial Monitoring	441
21.6	Analysis of Data	442
21.7	Interpretation of Outcomes	444

21.8	Regulatory Perspective	446
21.9	Summary and Discussion	447
22	Adaptive Dose-Escalation Trial	449
22.1	Oncology Dose-Escalation Trial	449
22.1.1	Dose Level Selection	449
22.1.2	Traditional Escalation Rules	450
22.1.3	Simulations Using Traditional Escalation Algorithms	453
22.2	Continual Reassessment Method	455
22.2.1	Probability Model for Dose-Response	456
22.2.2	Prior Distribution of Parameter	456
22.2.3	Reassessment of Parameter	457
22.2.4	Assignment of Next Patient	458
22.2.5	Stopping Rule	458
22.3	Alternative Form CRM	458
22.4	Simulations of CRM	459
22.5	Bayesian Model Averaging CRM	464
22.6	Summary and Discussion	465
23	Bayesian Design for Efficacy–Toxicity Trade-Off and Drug Combination	467
23.1	Introduction	467
23.2	Thall–Russell Independent Model	467
23.3	Efficacy–Toxicity Trade-Off Model	469
23.4	Odds Ratio Trade-off Model	470
23.5	Drug Combination	472
23.6	Summary	474
24	Bayesian Approach to Biosimilarity Trial	477
24.1	Introduction	477
24.2	Dilemma of Frequentist Noninferiority Trial	478
24.3	Synergic Constancy Approach	479
24.3.1	Conditional Bias of Historical Trial Data	479
24.3.2	Requirements for Biosimilarity or Noninferiority	482
24.4	Bayesian Approach Combining Preclinical and Clinical Data	487
24.5	Bayesian Hierarchical Bias Model	490
24.5.1	Biosimilarity Using Composite Endpoint	490

24.5.2	Study Design and Noninferiority Hypotheses . . .	491
24.5.3	Bayesian Approach	492
24.5.4	Hierarchical Bias Model	492
24.6	Summary	497
25	Adaptive Multiregional Trial Design	499
25.1	Introduction	499
25.2	Unified Additional Requirement for Regional Approval in MRCT	501
25.2.1	Current Consistency Study	501
25.2.2	The Unified Regional Requirement	504
25.2.3	Determination of Parameter λ and α_i	507
25.3	Optimal Classical MRCT Design	512
25.3.1	Maximum Power Design	512
25.3.2	Maximum Utility Design	513
25.4	Optimal Adaptive MRCT Design	513
25.5	Bayesian Approach	518
25.6	Practical Issues and Challenges	520
25.7	Summary	521
26	SAS and R Modules for Group Sequential Design	525
26.1	Introduction	525
26.2	SEQDESIGN Procedure	526
26.2.1	PROC SEQDESIGN Statement	526
26.2.2	DESIGN Statement	527
26.2.3	SAMPLESIZE Statement	527
26.3	Examples with Proc SEQDESIGN	530
26.4	SAS SEQTEST and R sgDesign	531
27	Data Analysis of Adaptive Trial	535
27.1	Introduction	535
27.2	p -value Calculation	536
27.3	Parameter Estimation	537
27.4	Confidence Interval	540
27.5	Parameter Estimation in Trials with Arm Selection	542
27.5.1	Parametric Bootstrap Approach	544
27.5.2	Shrinkage Estimators	547
27.5.3	Empirical Bayes Estimators	548
27.6	Summary	550

28	Planning, Execution, Analysis, and Reporting	553
28.1	Validity and Integrity	553
28.2	Study Planning	554
28.3	Working with Regulatory Agency	555
28.4	Trial Monitoring	559
28.5	Analysis and Reporting	560
28.6	Clinical Trial Simulation	561
28.7	Summary	562
29	Debates in Adaptive Designs	565
29.1	My Standing Point	565
29.2	Decision Theory Basics	566
29.3	Evidence Measure	567
29.3.1	Frequentist p -value	568
29.3.2	Bayes Factor	568
29.3.3	Bayesian p -value	569
29.3.4	Repeated Looks	570
29.3.5	Role of Alpha in Drug Development	570
29.4	Statistical Principles	571
29.5	Behaviors of Statistical Principles in Adaptive Designs . .	575
29.5.1	Sufficiency Principle	575
29.5.2	Minimum Sufficiency Principle and Efficiency . .	576
29.5.3	Conditionality and Exchangeability Principles . .	577
29.5.4	Equal Weight Principle	578
29.5.5	Consistency of Trial Results	579
29.5.6	Bayesian Aspects	580
29.5.7	Type-I Error, p -value, Estimation	581
29.5.8	The 0-2-4 Paradox	582
29.6	Summary	583
Appendix A	Random Number Generation	585
A.1	Random Number	585
A.2	Uniformly Distributed Random Number	585
A.3	Inverse CDF Method	586
A.4	Acceptance-Rejection Methods	586
A.5	Multivariate Distribution	587
Appendix B	A Useful Utility	591

Appendix C	SAS Macros for Add-Arm Designs	593
C.1	The 5+1 Add-Arm Design	593
C.2	The 6+1 Add-Arm Design	594
C.3	The 7+1 Add-Arm Design	596
Appendix D	Implementing Adaptive Designs in R	599
Bibliography		611

List of Figures

1.1	Trends in NDAs Submitted to FDA	2
1.2	Sample-Size Reestimation Design	4
1.3	Drop-Loser Design	5
1.4	Response Adaptive Randomization	6
1.5	Dose Escalation for Maximum Tolerated Dose	7
1.6	Biomarker-Adaptive Design	8
1.7	Adaptive Treatment Switching	9
1.8	Clinical Trial Simulation Model	11
1.9	Characteristics of Adaptive Designs	12
2.1	A Simplified View of the NDA	24
2.2	Power as a Function of α and n	26
2.3	Power and Probability of Efficacy (PE)	30
2.4	p -value versus Observed Effect Size	31
3.1	Bayesian Decision Approach	67
5.1	Examples of Brownian motion	115
8.1	Error-Spending Functions	160
9.1	Various Stopping Boundaries at Stage 2	191
10.1	Recursive Two-Stage Adaptive Design	204
11.1	Conditional Power versus p -value from Stage 1	233
12.1	Mixed Distribution Changes as the Means Change	246
14.1	Multiple-Endpoint Adaptive Design	301

15.1	Seamless Design	308
15.2	Decision Theory for Competing Constraints	309
16.1	The 4 + 1 Add-Arm Design	324
16.2	The 5 + 1 Add-Arm Design	333
16.3	The 6 + 1 Add-Arm Design	335
16.4	The 7 + 1 Add-Arm Design	336
17.1	Effect of Biomarker Misclassification	355
17.2	Treatment-Biomarker-Endpoint Three-Way Relationship	363
17.3	Correlation versus Prediction	363
18.1	Relationships between Biomarker and Primary Endpoint	374
19.1	A Simple Multistage Survival Model	401
19.2	Effect of Treatment Switching	403
20.1	Randomized Play-the-Winner	412
20.2	Sequential Parallel Design with Rerandomization	425
21.1	Bayesian Learning Process	429
21.2	ExpDesign Studio for Classical and Adaptive Designs . .	439
21.3	Interpretation of Confidence Interval: Five out of 100 in- tervals do not cover the population mean (0) with $\alpha =$ 5%	445
22.1	Logistic Toxicity Model	456
23.1	Efficacy–Toxicity Trade-Off Contours for Pentostatin Trail	470
23.2	Toxicity–Efficacy Odds Ratio Trade-off Contours	471
25.1	Partial Data from a Multiregional Trial	500
28.1	Simplified CTS Model: Gray-Box	562

List of Tables

2.1	Sample Sizes for Different Types of Endpoints	27
2.2	Sample Sizes for Different Contrasts (Balanced Design) .	44
2.3	Response and Contrast Shapes	44
2.4	Sample Size per Group for Various Contrasts	45
3.1	Conditional and Unconditional Means	61
4.1	Stopping Boundaries α_2 with MIP	72
4.2	Operating Characteristics of a GSD with MIP	75
4.3	Stopping Boundaries with MSP	78
4.4	Stopping Boundaries α_2 with MSP	78
4.5	Operating Characteristics of a GSD with MSP	81
4.6	Stopping Boundaries α_2 with MPP	83
4.7	Operating Characteristics of a GSD with MPP	87
4.8	Operating Characteristics of a GSD with MSP	89
4.9	Operating Characteristics of a GSD with MSP	91
4.10	Operating Characteristics of a GSD with MSP	92
4.11	Operating Characteristics of Adaptive Methods	94
5.1	Stopping Boundaries α_2 with Equal Weights	105
5.2	Operating Characteristics of GSD with MINP	109
5.3	Operating Characteristics of an SSR with MSP	111
5.4	Operating Characteristics of a GSD with MSP	114
6.1	Matched-Pair Data	126
6.2	Type-I Error Rate Control (%) against $\alpha = 2.5\%$	130
6.3	Sensitivity and Specificity	130
6.4	Positive Patients per CT/Bone Scan	131

6.5	Negative Patients per CT/Bone Scan	131
6.6	Operating Characteristics of AD under H_a for Sensitivity	133
6.7	Operating Characteristics of Adaptive Design for Specificity	134
6.8	Power Preserved by GSD and SSR Designs for Specificity	134
7.1	Normal Endpoint (Sample Size $n = 500$)	147
7.2	Binary Endpoint (Sample Size $n = 500$)	147
7.3	Binary Endpoint (Sample Size $n = 100$)	147
8.1	Stopping Boundaries of Three-Stage Design with MSP	159
8.2	3-Stage Design Operating Characteristics without SSR	168
8.3	3-Stage Design Operating Characteristics with SSR	169
8.4	4-Stage Design Operating Characteristics without SSR	173
8.5	4-Stage Design Operating Characteristics with SSR	174
8.6	Two-Arm Design Operating Characteristics without Adjustment	178
8.7	Two-Arm Operating Characteristics with SSR	179
9.1	Stopping Boundaries with MMP ($w_1 = w_2 = 1/2$)	189
9.2	Stopping Boundaries with MLP ($w_1 = 1/3, w_2 = 2/3$)	189
9.3	Example Stopping Boundaries with MPPP ($v = 1.5$)	190
9.4	Conditional Error Functions (under H_0)	190
9.5	Stopping Boundaries without Futility Binding	190
9.6	B -Function for Conditional Power and Sample-Size Reestimation	193
9.7	Comparisons of Conditional Powers cP	195
10.1	Summary of the Recursive Two-Stage Design	209
11.1	Comparisons of Adaptive Design Methods ($\alpha_1 = 0.0025$)	222
11.2	Comparisons of Adaptive Design Methods ($\alpha_1 = 0.01$)	222
11.3	Comparisons of Adaptive Design Methods with ($\beta_1 = 0.15$)	222
11.4	Comparisons of Adaptive Methods	231
11.5	Summary of Comparisons with Lower Initial Power	231
11.6	Conditional Power as Function of N_2	232
11.7	Conditional Powers as Function of p_1	233
11.8	Comparison of Adaptive Designs	238

12.1	Sample Size per Group Required ($\sigma - \kappa$ Method)	250
12.2	Sample Size per Group Required (Semi-Blinded Method)	252
12.3	Comparisons of SSR Methods	254
12.4	Comparisons of SSR Mixed Methods	259
13.1	Overall Power	272
13.2	Maximum Sample Size with Overall Power 80%	273
13.3	Overall Conditional Power	275
14.1	Error Inflation Due to Correlations between Endpoints	285
14.2	Error Inflation Due to Different Numbers of Endpoints	286
14.3	MINP Based on Hypothetical p_{ik}	303
14.4	RTAD Based on Hypothetical p_{ik}	305
15.1	Critical Value c_α for Classical Pick-the-Winner Design	313
16.1	c_R, c_α and Percent of α -spent ($\alpha = 0.025, \tau = 0.5$)	339
16.2	Responses for Difference Response Curves	340
16.3	Power (%) Comparisons	341
16.4	Comparison of Power and Selection Probability	341
16.5	Phase II Dose-Finding Models	343
16.6	Power and Selection Probability: E_{max} Model	344
16.7	Power and Selection Probability: Linear in Log-Dose	344
16.8	Power and Selection Probability: Linear	344
16.9	Power and Selection Probability: Truncated-Logistic	344
16.10	Power and Selection Probability: Logistic	344
16.11	Selection Probability and Sample Size: AA and DA Designs	345
17.1	Response Rate and Sample Size Required	353
17.2	Simulation Results of Two-Stage Design	360
17.3	Issues with Biomarker Primary Endpoint	362
17.4	Adaptive Design with Biomarker	367
17.5	Prior Knowledge about Effect Size	367
17.6	Expected Utilities of Different Designs	367
18.1	Power of Winner Design with One-Level Correlation	374
18.2	Critical Value of Two-Stage Winner Design	379
18.3	Power of Biomarker-Informed Design	381
18.4	Biomarker-Informed Design with MEM (ρ and σ Effects)	385
18.5	Biomarker-Informed Design with MEM (μ Effect)	385

20.1	Asymptotic Variance with RPW	414
20.2	Simulation Results from RPW	418
21.1	Commonly Used Conjugate Families	431
21.2	Estimation of Conjugate Families	432
21.3	Prior Knowledge	438
21.4	Characteristics of Classical Phase II and III Designs . . .	438
21.5	Characteristics of Seamless Design (OF)	440
21.6	Characteristics of Seamless Design (Pocock)	440
21.7	Comparison of Classical and Seamless Designs	441
22.1	Dose Levels and DLT Rates	462
22.2	Adaptive Dose-Response Simulation Results	463
22.3	Dose Levels and DLT Rates	466
24.1	Dilemma of Noninferiority Trial	479
24.2	Relationship between Bias ($\times(\mu_1 - \mu_0)/\sigma$) and Power . .	481
25.1	An Example of Global Multiregional Clinical Trial	500
25.2	Sample Size Ratio R for Various AP (AP_i^0)	507
25.3	λ Values for Various Sample Size Ratios	508
25.4	α_i Values for Various Sample Size Ratios	508
25.5	α_i Values for Various λ and Sample Size Ratios	508
25.6	Conditional and Unconditional Error Rates	509
25.7	Comparison of Classical and Adaptive MRCT Designs . .	518
26.1	Main Options in Proc SEQDESIGN	526
26.2	Main Options in DESIGN Statement	527
26.3	Main Options in SAMPLESIZE Statement	528

List of Examples

Example 2.1	Arteriosclerotic Vascular Disease Trial	29
Example 2.2	Equivalence LDL Trial	32
Example 2.3	Average Bioequivalence Trial	36
Example 2.4	Dose-Response Trial with Continuous Endpoint .	41
Example 2.5	Dose-Response Trial with Binary Endpoint	42
Example 2.6	Dose-Response Trial with Survival Endpoint . . .	43
Example 4.1	Adaptive Design for Acute Ischemic Stroke Trial .	74
Example 4.2	Adaptive Design for Asthma Study	80
Example 4.3	Adaptive Design for Oncology Trial	85
Example 4.4	Early Futility Stopping Design with Binary Endpoint	88
Example 4.5	Noninferiority Design with Binary Endpoint . . .	89
Example 4.6	Sample-Size Reestimation with Normal Endpoint	90
Example 4.7	Sample-Size Reestimation with Survival Endpoint	93
Example 4.8	Adaptive Equivalence LDL Trial	99
Example 5.1	Inverse-Normal Method with Normal Endpoint .	108
Example 5.2	Inverse-Normal Method with SSR	110
Example 5.3	Group Sequential Design	112
Example 5.4	Changes in Number and Timing of Interim Analyses	119
Example 7.1	Superiority Trial Designs with Normal Endpoint .	149
Example 7.2	Noninferiority Trial Designs with Normal Endpoint	149
Example 7.3	Equivalence Trial Design with Normal Endpoint .	150
Example 7.4	Superiority Trial Design with Binary Endpoint . .	151
Example 7.5	A Clinical Trial of Retinal Diseases	152
Example 8.1	Three-Stage Adaptive Design	166
Example 8.2	Four-Stage Adaptive Design	172
Example 8.3	Adaptive Design with Survival Endpoint	176
Example 9.1	Adaptive Design for Coronary Heart Disease Trial	185

Example 10.1	Recursive Two-Stage Adaptive Design	206
Example 10.2	Application of Recursive Combination Method . .	211
Example 11.1	Myocardial Infarction Prevention Trial	236
Example 11.2	Adaptive Design with Farrington-Manning NI Margin	239
Example 12.1	Blinded Sample-Size Reestimation for Binary Endpoint	264
Example 13.1	Power and Sample Size for Two-Arm Trial with Coprimary Endpoints	275
Example 13.2	Conditional Power and Sample-Size Reestimation in Single/Paired Group Trial with Coprimary Endpoints	277
Example 13.3	Conditional Power for Two-Group Design with Two Coprimary Endpoints	278
Example 14.1	A Union-Intersection Test	283
Example 14.2	Intersection-Union Test with Alzheimer's Trial . .	284
Example 14.3	Union-Intersection Mixture Testing	284
Example 14.4	Multiple Testing for Dose-Finding Trial	293
Example 14.5	Multiple Testing for Primary and Secondary Endpoints	297
Example 14.6	Three-Stage Adaptive Design for NHL Trial . . .	300
Example 15.1	Seamless Design of Asthma Trial	314
Example 16.1	Phase II Dose-Finding Trial	343
Example 16.2	Phase II-III Asthma Trial	346
Example 16.3	Phase IV Oncology Trial	347
Example 17.1	Biomarker-Adaptive Design	360
Example 19.1	The First-Hitting-Time Model for Biological Degradation	397
Example 19.2	Treatment Switch in Survival Trial	397
Example 19.3	Multistage Model for Survival Trial	402
Example 19.4	Multistage Model for Clinical Trial Treatment Switch	403
Example 20.1	Randomized Play-the-Winner Design	416
Example 20.2	Adaptive Randomization with Normal Endpoint .	422
Example 21.1	Beta Posterior Distribution	430
Example 21.2	Normal Posterior Distribution	430
Example 21.3	Prior Effect on Power	432
Example 21.4	Power with Normal Prior	433
Example 21.5	Bayesian Power	434

Example 21.6	Trial Design Using Bayesian Power	435
Example 21.7	Simon Two-Stage Optimal Design	437
Example 21.8	Bayesian Optimal Design	438
Example 22.1	Adaptive Dose-Finding for Prostate Cancer Trial	462
Example 24.1	Biosimilar Diabetic Trial Design	486
Example 25.1	Multiregional ACS Clinical Trial	511
Example 25.2	Adaptive Multiregional ACS Clinical Trial	517
Example 29.1	Paradox of Binomial and Negative Binomial Distribution	574

List of SAS Macros and R Functions

SAS Macro 2.1:	Equivalence Trial with Normal Endpoint . . .	32
SAS Macro 2.2:	Equivalence Trial with Binary Endpoint . . .	33
SAS Macro 2.3:	Crossover Bioequivalence Trial	36
SAS Macro 2.4:	Sample Size for Dose-Response Trial	46
SAS Macro 4.1:	Two-Stage Adaptive Design with Binary Endpoint	73
SAS Macro 4.2:	Two-Stage Adaptive Design with Normal Endpoint	79
SAS Macro 4.3:	Two-Stage Adaptive Design with Survival Endpoint	84
SAS Macro 4.4:	Event-Based Adaptive Design	95
SAS Macro 4.5:	Adaptive Equivalence Trial Design	97
SAS Macro 5.1:	Stopping Boundaries with Adaptive Designs	104
SAS Macro 5.2:	Two-Stage Design with Inverse-Normal Method	107
SAS Macro 6.1:	Adaptive Noninferiority Design with Paired Data	135
SAS Macro 7.1:	Sample-Size Reestimation Design with Incomplete Pairs	153
SAS Macro 8.1:	K-Stage Adaptive Designs with Normal Endpoint	165
SAS Macro 8.2:	K-Stage Adaptive Designs with Binary Endpoint	170
SAS Macro 8.3:	K-Stage Adaptive Designs with Various Endpoints	175
SAS Macro 9.1:	Conditional Power	193
SAS Macro 9.2:	Sample Size Based on Conditional Power . . .	195

SAS Macro 11.1:	Two-Stage Sample-Size Reestimation	223
SAS Macro 11.2:	K -Stage Adaptive Design with Sample-Size Reestimation	226
SAS Macro 12.1:	Blinded $\sigma - K$ Method for Mean Difference .	250
SAS Macro 12.2:	Semi-blinded Method for Mean Difference . .	252
SAS Macro 12.3:	Blinded and Unblinded Sample Size Reestimation	256
SAS Macro 12.4:	Sample-Size Reestimation with Mixed Methods	258
SAS Macro 12.5:	Sample-Size Reestimation with Shih-Zhao Method	263
R Function 13.1:	Power of Two Coprimary Endpoints	273
R Function 13.2:	Power of Two Coprimary Endpoints by Simulation	274
R Function 13.3:	Overall Conditional Power for One-Group Design with Two Coprimary Endpoints	278
R Function 13.4:	Overall Conditional Power for Two-Group Design with Two Coprimary Endpoints	279
SAS Macro 15.1:	Pick-the-Winner Design	313
SAS Macro 16.1:	4+1 Add-Arm Design	329
SAS Macro 16.2:	4+1 Add-Arm Design for Finding MED . . .	330
SAS Macro 17.1:	Biomarker-Adaptive Design	359
R Function 18.1:	Stopping Boundary of Biomarker-Informed Design	379
R Function 18.2:	Biomarker-Informed Design with Hierarchical Model	383
SAS Macro 20.1:	Randomized Play-the-Winner Design	415
SAS Macro 20.2:	Binary Response-Adaptive Randomization . .	419
SAS Macro 20.3:	Normal Response-Adaptive Randomization . .	421
SAS Macro 21.1:	Simon Two-Stage Futility Design	436
SAS Macro 22.1:	3+3 Dose-Escalation Design	454
SAS Macro 22.2:	Continual Reassessment Method	460
SAS Macro 24.1:	Publication Bias	481
SAS Macro 24.2:	Biosimilar Clinical Trial	485
SAS Macro 25.1:	Classical Multireginal Clinical Trial	510
SAS Macro 25.2:	Adaptive Multiregional Clinical Trial	515
SAS Macro A.1:	Mixed Exponential Distribution	587
SAS Macro A.2:	Multivariate Normal Distribution	589
SAS Macro C.3:	The 5+1 Add-Arm Design	593

SAS Macro C.4:	The 6+1 Add-Arm Design	594
SAS Macro C.5:	The 7+1 Add-Arm Design	596
R Function D.1:	Sample-Size Based on Conditional Power . . .	599
R Function D.2:	Sample-Size Reestimation	601
R Function D.3:	Biomarker-Adaptive Design	603
R Function D.4:	Randomized Play-the-Winner Design	605
R Function D.5:	Continual Reassessment Method	607

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,

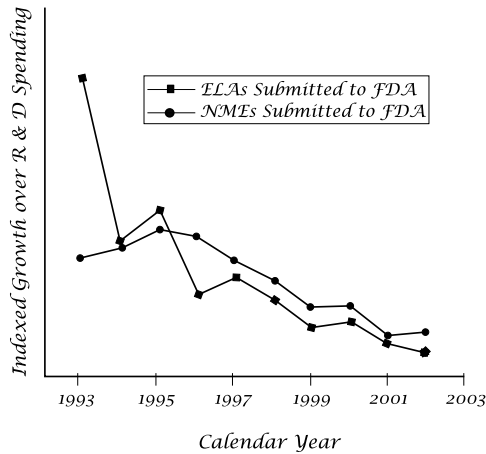


Figure 1.1: Trends in NDAs Submitted to FDA (Data Source: PAREXEXL, 2003)

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) a 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 (EMA, 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.

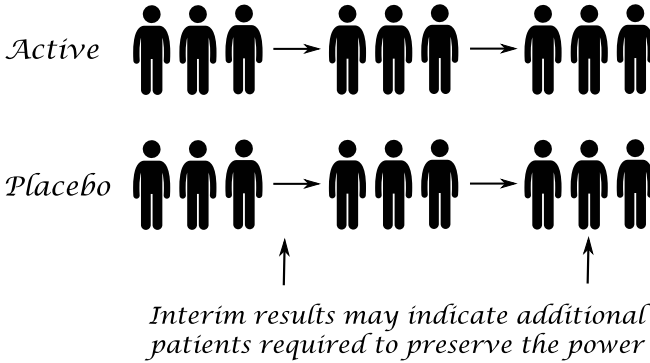


Figure 1.2: Sample-Size Reestimation Design

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.

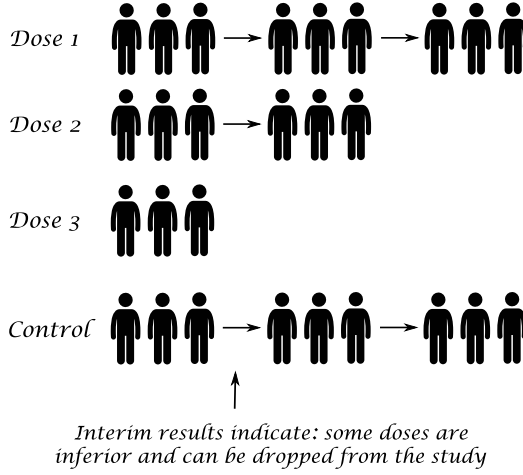


Figure 1.3: Drop-Loser Design

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).

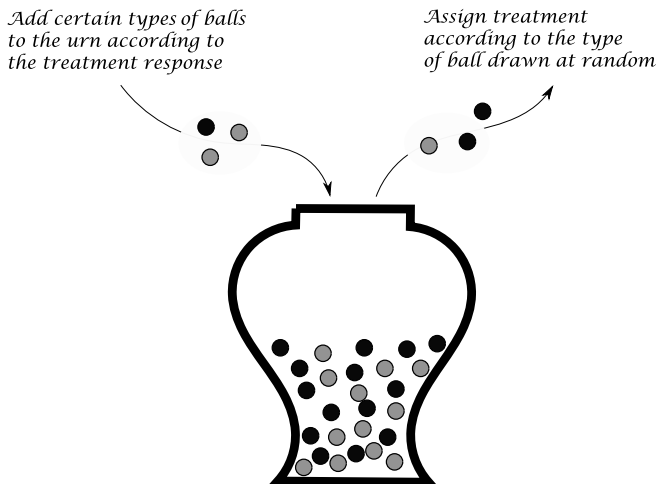


Figure 1.4: Response Adaptive Randomization

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.

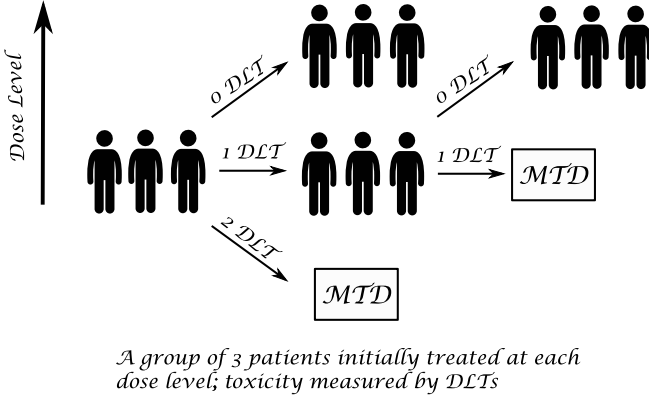


Figure 1.5: Dose Escalation for Maximum Tolerated Dose

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).

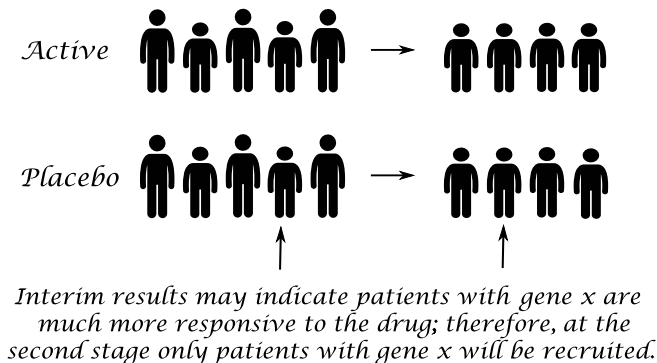


Figure 1.6: Biomarker-Adaptive Design

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).

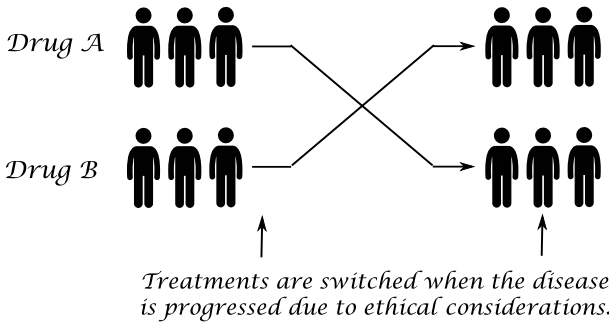


Figure 1.7: Adaptive Treatment Switching

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

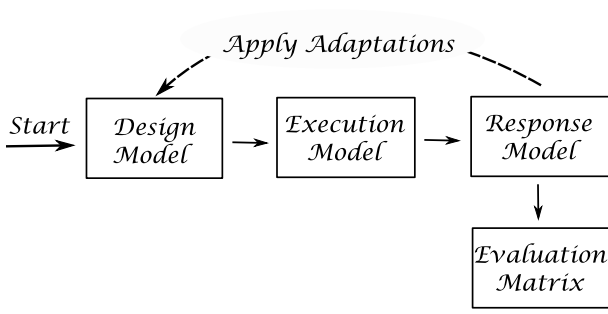


Figure 1.8: Clinical Trial Simulation Model

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.

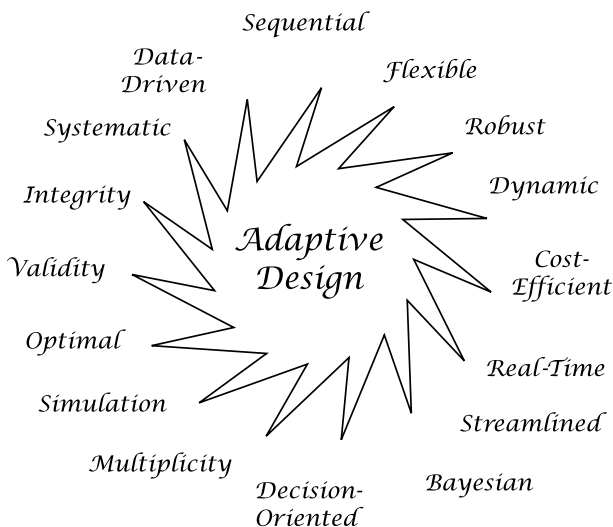


Figure 1.9: Characteristics of Adaptive Designs

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 (EMA)—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 stage-wise 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/allocation 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 post-marketing surveillance trials, other reasons sponsors may conduct post-marketing 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).

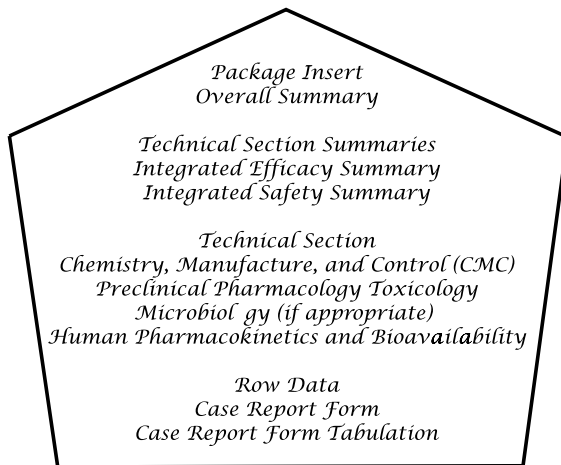


Figure 2.1: A Simplified View of the NDA

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