Biomarker informed adaptive clinical trial designs
BIOMARKER INFORMED ADAPTIVE CLINICAL TRIAL DESIGNS

by

JING WANG
B.S., University of Science and Technology of China, 2008
M.S., University of Kentucky, 2010

Submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
2014
DEDICATION

To the Memory of My Beloved Grandparents.
ACKNOWLEDGMENTS

I would like to extend my sincere thanks to Dr. Mark Chang, who has been always there to help me to learn and to keep me motivated throughout my journey. I greatly appreciate his support and care. I would never have made this far without his mentoring.

I would also like to thank my dissertation committee: Dr. Gheorghe Doros, Dr. Joseph Massaro, Dr. Sandeep Menon and the committee chair Dr. Michael LaValley for their help and expertise.

The journey of graduate study could never have been so joyful without the encouragement and support from my supervisors and friends. Special thanks to my academic advisor Dr. Qiong Yang and my internship managers Dr. Ling Wang and Dr. Bin Zhang. Thank you for offering me the opportunities to participate in the interesting research projects, and thank you for all your instructions and help! It had been great pleasure working with you! Thanks to the professors in the department for all the help and advice through my entire graduate work. Thanks to my friends: Jacqui, Susan, Joseph, Seung-hoan, Baiyun, Flora, Zhaoyang, Wei, thank you for all your help and the great time we shared!

Lastly and most importantly, I thank my family for their constant support and unconditional love!
ABSTRACT

In adaptive design clinical trials, an endpoint at the final analysis that takes a long time to observe is not feasible to be used for making decisions at the interim analysis. For example, overall survival (OS) in oncology trials usually cannot be used to make interim decisions. However, biomarkers correlated to the final clinical endpoint can be used. Hence, considerable interest has been drawn towards the biomarker informed adaptive clinical trial designs.

Shun et al. (2008) proposed a “biomarker informed two-stage winner design” with 2 active treatment arms and a control arm, and proposed a normal approximation method to preserve type I error. However, their method cannot be extended to designs with more than 2 active treatment arms. In this dissertation, we propose a novel statistical approach for biomarker informed two-stage winner design that can accommodate multiple active arms and control type I error. We further propose another biomarker informed adaptive design called “biomarker informed add-arm design for unimodal response”. This design utilizes existing knowledge about the shape of dose-response relationship to optimize the procedure of selecting best candidate treatment for a larger trial. The key element of the proposed design is that, some inferior treatments do not need to be explored and the
design is shown to be more efficient than biomarker informed two-stage winner design mathematically.

Another important component in the study of biomarker informed adaptive designs is to model the relationship between the two endpoints. The conventional approach uses a one-level correlation model, which might be inappropriate if there is no solid historical knowledge of the two endpoints. A two-level correlation model is developed in this dissertation. In the new model a new variable that describes the mean level correlation is developed, so that the uncertainty of the historical knowledge could be more accurately reflected. We use this new model to study the “biomarker informed two-stage winner design” and the “biomarker informed add-arm design for unimodal response”. We show the new proposed model performs better than conventional model via simulations.

The concordance of inference based on biomarker and primary endpoint is further studied in a real case.
TABLE OF CONTENTS

DEDICATION iv
ACKNOWLEDGMENTS v
ABSTRACT vi
TABLE OF CONTENTS viii
LIST OF TABLES xii
LIST OF FIGURES xiv
LIST OF ABBREVIATIONS xv

CHAPTER 1 INTRODUCTION

1.1 Conventional Adaptive Clinical Trial Designs ............................................................ 1
  1.1.1 Group Sequential Design ......................................................................................... 3
  1.1.2 Sample-Size Re-estimation Design ........................................................................ 4
  1.1.3 Drop-the-Loser Design ......................................................................................... 6
  1.1.4 Add-Arm Design for Unimodal Response .............................................................. 7
  1.1.5 Other Designs ...................................................................................................... 8

1.2 Biomarker Informed Adaptive Clinical Trial Designs ............................................. 9
  1.2.1 Biomarker Informed Two-Stage Winner Design ................................................... 11
  1.2.2 Biomarker Informed Add-Arm Design for Unimodal Response ......................... 12

1.3 Structure of the Dissertation .................................................................................... 13

CHAPTER 2 FINDING CRITICAL VALUES WITH CONTROLLED TYPE I
ERROR FOR A BIOMARKER INFORMED TWO-STAGE WINNER DESIGN

2.1 Motivation ................................................................................................................ 15
2.2 Finding Critical Rejection Values for Final Test Statistic in Biomarker Informed Two-Stage Winner Design ......................................................... 19
  2.2.1 Biomarker Informed Two-Stage Winner Design .................................. 19
  2.2.2 The Two Models for Fitting the Two Endpoints ..................................... 20
  2.2.3 Test Statistic and its Distribution ......................................................... 23
  2.2.4 Critical Rejection Values and $R$ .......................................................... 29
  2.3 Discussion .................................................................................................. 35

CHAPTER 3 MODELING THE BIOMARKER INFORMED ADAPTIVE CLINICAL TRIAL DESIGNS

  3.1 Introduction ................................................................................................. 36
  3.2 Motivations and Concepts ........................................................................... 38
  3.3 Issues in Conventional One-Level Correlation Model ................................. 40
  3.4 Biomarker Informed Two-Stage Winner Design with the Proposed Two-Level Correlation Model ................................................................. 44
    3.4.1 The Biomarker Informed Two-Stage Winner Design ............................... 44
    3.4.2 The Proposed Two-Level Correlation Model ......................................... 45
    3.4.3 Test Statistic and its Distribution .......................................................... 47
    3.4.4 Type I Error Control ............................................................................ 49
  3.5 Performance Evaluation of Two-Stage Winner Design under Proposed Model .... 51
    3.5.1 $u^x$ and $u^y$ are Linearly Related ...................................................... 52
    3.5.2 $u^x$ and $u^y$ are Not Linearly Related ................................................ 54
  3.6 Parameter Estimation ................................................................................... 58
CHAPTER 4  BIOMARKER INFORMED ADD-ARM DESIGN FOR UNIMODAL RESPONSE

4.1 Introduction.................................................................................................................. 63

4.2 Biomarker Informed Add-Arm Design with 4 Active Treatment Arms .............. 66
  4.2.1 The 3-Stage Design.................................................................................................. 66
  4.2.2 Mathematical Notations for the Design ................................................................. 68
  4.2.3 Find $c_R$ ............................................................................................................... 71
  4.2.4 Type I Error Control ............................................................................................. 73
  4.2.5 Power .................................................................................................................... 78

4.3 Biomarker Informed Add-Arm Design with $K$ ($K > 4$) Active Treatments .... 86
  4.3.1 Biomarker Informed Add-Arm Design with 5 Active Treatments ................... 86
  4.3.2 Biomarker Informed Add-Arm Design with 6 Active Treatments ............... 89
  4.3.3 Biomarker Informed Add-Arm Design with 7 Active Treatments ............... 91

4.4 Discussion.................................................................................................................. 94

CHAPTER 5  A CASE STUDY OF USING BIOMARKER AT INTERIM

5.1 Introduction.................................................................................................................. 96

5.2 Survival Model in Non-Small-Cell Lung Cancer ..................................................... 96

5.3 Data Generation Based On Survival Model for Non-Small-Cell Lung Cancer .... 98

5.4 Inference Based Upon Change in Tumor Size vs. Inference Based On Survival... 99

5.5 Discussion.................................................................................................................. 105
CHAPTER 6  SUMMARY AND USING A BIOMARKER INFORMED ADAPTIVE DESIGN IN AN A&WC STUDY ......................................................... 107

APPENDIX A  R Functions for Calculation of Critical Rejection Values for Biomarker Informed Two-Stage Winner Design and Biomarker Informed Add-Arm Design for Unimodal Response 111

APPENDIX B  Derivation for $\hat{\sigma}^2_j$ 115

APPENDIX C  Proof of the Equations 116

REFERENCES 117

CURRICULUM VITAE 125
LIST OF TABLES

Table 2.1 \( w_{0.025} \) for biomarker informed two-stage winner design \((K = 4)\) .................................. 33
Table 2.2 \( w_{0.025} \) for biomarker informed two-stage winner design \((K = 5)\) .................... 33
Table 2.3 \( w_{0.025} \) for biomarker informed two-stage winner design \((K = 6)\) .................... 33
Table 2.4 \( w_{0.025} \) for biomarker informed two-stage winner design \((K = 7)\) ................ 34
Table 2.5 \( w_{0.025} \) for biomarker informed two-stage winner design \((K = 8)\) ................. 34

Table 3.1 Power Evaluation Using Conventional One-Level Correlation Model ............. 42
Table 3.2 Power when \( u^X \) and \( u^Y \) are Linearly Related \((n_1 = 52, N = 104)\) ........ 53
Table 3.3 Power when \( u^X \) and \( u^Y \) are Linearly Related \((n_1 = 67, N = 134)\) ........ 53
Table 3.4 Power when \( u^X \) and \( u^Y \) are Not Linearly Related – Case 1 \((n_1 = 52, N =
104)\) .................................................................................................................. 55

Table 3.5 Power when \( u^X \) and \( u^Y \) are Not Linearly Related – Case 2 \((n_1 = 52, N =
104)\) .................................................................................................................. 55

Table 3.6 Power with Best Treatment when \( u^X \) and \( u^Y \) are Linearly Related \((n_1 =
52, N = 104)\) .................................................................................................. 56
Table 3.7 Power with Best Treatment when \( u^X \) and \( u^Y \) are Linearly Related \((n_1 =
67, N = 134)\) .................................................................................................. 56
Table 3.8 Power with Best Treatment when \( u^X \) and \( u^Y \) are Not Linearly Related – Case
1 \((n_1 = 52, N = 104)\) ...................................................................................... 57
Table 3.9 Power with Best Treatment when \( u^X \) and \( u^Y \) are Not Linearly Related – Case
2 \((n_1 = 52, N = 104)\) ...................................................................................... 57
Table 3.10 Estimation of the Parameters ............................................................................. 60
Table 4.1 Critical Rejection Value w_{0.025} ................................................................. 78
Table 4.2 Power: Biomarker Informed Two-Stage Winner Design (Linear) .................. 82
Table 4.3 Power: Biomarker Informed Two-Stage Winner Design (Nonlinear) ........... 82
Table 4.4 Power: Biomarker Informed Add-Arm Design (Linear) .............................. 83
Table 4.5 Power: Biomarker Informed Add-Arm Design (Nonlinear) ........................ 84
Table 5.1 Concordance Rate: Design Specifications Assumed Correctly ................. 101
Table 5.2 Concordance Rate: $\sigma_T^2$ Assumed Incorrectly .......................................... 102
Table 5.3 Concordance Rate: $u_{T_1}$ Assumed Incorrectly ............................................. 103
LIST OF FIGURES

Figure 1.1 Pharma R&D Spending: 1995-2009................................................................. 1
Figure 1.2 Two-Stage Winner Design ........................................................................ 7
Figure 1.3 Biomarker Informed Two-Stage Winner Design ........................................ 12
Figure 3.1 Illustration of Relationship Between Biomarker and Primary Endpoint ...... 43
Figure 3.2 Mean Level Correlation Shapes .................................................................. 52
Figure 4.1 Biomarker Informed Add-Arm Design (4+1 Case)........................................ 67
Figure 4.2 Detailed Biomarker Informed Add-Arm Design (4+1 Case)......................... 70
Figure 4.3 Mean Level Correlation Shapes ................................................................. 80
Figure 4.4 Dose Response Curves ............................................................................. 81
Figure 4.5 Biomarker Informed Add-Arm Design for Unimodal Response (5+1 Case) .. 86
Figure 4.6 Biomarker Informed Add-Arm Design for Unimodal Response (6+1 Case) .. 89
Figure 4.7 Biomarker Informed Add-Arm Design for Unimodal Response (7+1 Case) .. 91
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARD</td>
<td>Adaptive Randomization Design</td>
</tr>
<tr>
<td>ATSD</td>
<td>Adaptive Treatment-Switching Design</td>
</tr>
<tr>
<td>A&amp;WC</td>
<td>Adequate and Well-Controlled</td>
</tr>
<tr>
<td>CRM</td>
<td>Continual Reassessment Method</td>
</tr>
<tr>
<td>FDA</td>
<td>Food and Drug Administration</td>
</tr>
<tr>
<td>GAO</td>
<td>Government Accountability Office</td>
</tr>
<tr>
<td>GSD</td>
<td>Group Sequential Design</td>
</tr>
<tr>
<td>MTD</td>
<td>Maximum Tolerated Dose</td>
</tr>
<tr>
<td>NDA</td>
<td>New Drug Application</td>
</tr>
<tr>
<td>ORR</td>
<td>Overall Response Rate</td>
</tr>
<tr>
<td>OS</td>
<td>Overall Survival</td>
</tr>
<tr>
<td>PFS</td>
<td>Progression Free Survival</td>
</tr>
<tr>
<td>RAR</td>
<td>Response-Adaptive Randomization</td>
</tr>
<tr>
<td>SSR</td>
<td>Sample-Size Re-estimation</td>
</tr>
<tr>
<td>TD</td>
<td>Time to Death</td>
</tr>
<tr>
<td>TSC</td>
<td>Tumor Size Change</td>
</tr>
<tr>
<td>UMPCU</td>
<td>Uniformly Most Powerful Conditionally Unbiased</td>
</tr>
</tbody>
</table>
CHAPTER 1 INTRODUCTION

1.1 Conventional Adaptive Clinical Trial Designs

The cost of drug development has increased dramatically in recent years. According to the NIH Office of Budget (Figure 1.1), the investment in pharmaceutical research and development has more than doubled in the past decade. However, the success rate for new drug applications (NDAs) remains low. As reported by the U.S. Government Accountability Office (GAO), the approval rate for NDAs submitted to the FDA in 2009 is only about 40%.

![Figure 1.1 Pharma R&D Spending: 1995-2009](image)

Reasons for this include (Woodcock, 2005):

- a diminished margin for improvement has escalated the level of difficulty in proving drug benefits;
- genomics and other new science have not yet reached their full potential;
mergers and other business arrangements have decreased candidates;

- easy targets are more difficult to study;

- rapidly escalating costs and complexity have decreased willingness to bring many candidates forward into the clinic.

To improve clinical development, great efforts have been made to develop innovative approaches, especially on adaptive designs, which allow adaptations or modifications to aspects of the trial after its initiation without undermining the validity and integrity of the trial (Chang, 2005). The PhRMA Working Group defines an adaptive design as a clinical study design that uses accumulating data to direct modification of aspects of the study as it continues, without undermining the validity and integrity of the trial (Gallo, et al., 2006).

Commonly employed adaptive designs in clinical trials use the same study endpoint at the interim and final stages of a trial. These designs are called “conventional adaptive clinical trial designs”, which include, but are not limited to:

- group sequential design;
- sample-size re-estimation design;
- drop-the-loser design;
- add-arm design for unimodal response;
- adaptive randomization design;
- adaptive dose-finding design;
- biomarker-adaptive design;
• adaptive treatment-switching design;
• Bayesian adaptive design.

1.1.1 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
(Chang, 2008). Basically, there are three different types of GSDs: early efficacy stopping
design, early futility stopping design, and early efficacy/futility stopping design.
The design typically leads to savings in sample size, time, and cost when compared with
a classical design with a fixed sample size. 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.

Different statistical methods have been proposed to avoid an inflation of the type I error
proposed critical values for the test statistic at each stage. These critical values were
made for two-sided tests. The one-sided versions were proposed by DeMets and Ware
(1980), DeMets and Ware (1982) and Pampallona and Tsiatis (1994). Lan and DeMets
(1983) significantly increased the flexibility of the group sequential concept by
introducing the alpha-spending function (or error spending function) approach, based on
the research of Slud and Wei (1982). Lan and DeMets (1983) presented closed form
spending functions that approximate the designs of Pocock (1977) and O’Brien and
Fleming (1979) in the case of equal stage sizes. More flexible families of spending
functions were proposed by Kim and DeMets (1987) and Hwang, Shih and DeCani (1990). Tsiatis and Robins (1997) developed a very general formulation of a group-sequential procedure based on the statistical information available at the times of the interim analyses. This “information-based design and monitoring procedure” can be applied to any type of model provided that there is a unique parameter of interest that can be efficiently tested. Applications were given, for example, by Mehta and Tsiatis (2001). In FDA guidance for industry regarding adaptive design clinical trials for drugs and biologics, group sequential design is categorized as a well-understood adaptive approach.

1.1.2 Sample-Size Re-estimation Design

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. 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 re-estimate the sample-size in the middle of the trial.

A sample-size re-estimation (SSR) design refers to an adaptive design that allows for sample-size adjustment or re-estimation based on the review of interim analysis results. There are two types of sample-size re-estimation procedures: sample-size re-estimation based on blinded data and sample-size re-estimation based on un-blinded data.

For sample-size re-estimation based on blinded data, the sample adjustment is based on the observed pooled variance at the interim analysis to recalculate the required sample
size, which does not require un-blinding the data. Type I error adjustment is practically negligible in this scenario. Gould and Shih (1992) and Friede and Kieser (2001) presented the methods. The sample-size re-estimation based on blinded data is categorized as a well-understood adaptive approach in FDA guidance for industry regarding adaptive design clinical trials for drugs and biologics.

For sample-size re-estimation based on un-blinded data, the effect size and its variability are re-assessed at interim stages, and sample size is adjusted based on the updated information. The type I error rate might be inflated in this scenario. Wittes, et al.(1999), Zucker, et al. (1999), Coffey and Muller (2001), and Denne and Jennison (1999) investigated internal pilot studies for this scenario with continuous data; Herson, et al. (1993), and Jennison, et al. (2003) discussed the scenario for binary data. Combination tests were proposed for Type I error control. Proschan and Hunsberger (1995) and Liu and Chi (2001) utilized conditional error functions; Lehmacher and Wassmer (1999) used inverse normal method; and Fisher (1998) introduced variance spending method. Chen, DeMets and Lan (2004) showed that if increasing sample size when the conditional power is greater than 50% at the interim, the regular unweighted test statistic can still be used without inflating the type I error rate. Gao, et al. (2008) used Brownian motion to derive a method of sample size re-estimation that achieves specified power against an alternative hypothesis corresponding to the current point estimate of the treatment effect.

The sample-size re-estimation based on un-blinded data is categorized as less well-understood adaptive approach in FDA guidance for industry regarding adaptive design clinical trials for drugs and biologics.
1.1.3 Drop-the-Loser Design

A drop-the-loser design is a design that allows dropping the inferior treatment groups (Chow and Chang, 2008). It is useful in phase II or phase II/III clinical development especially when there are uncertainties regarding dose levels. In drop-the-loser design, dose groups that are dropped may contain valuable information regarding dose response of the treatment under study. Therefore, the selection criteria and decision rules play important roles in the design. Ivanova (2003), Sun, et al (2006), Li, et al (2009), Wu and Zhao (2012) proposed drop-the-loser designs with different selection criteria.

Typically, a drop-the-loser design is a two-stage design with one best arm retained. This type of design is also called “two-stage winner design” (Figure 1.2). A two-stage winner design starts with several active treatment arms and a control arm. At interim, the inferior arms will be terminated based on pre-specified criteria, and only the most promising treatment (“winner”) will be carried until the end of the study. The final comparison between the winner arm and the control arm will be performed on data from both stages and on study primary endpoint. Notice that the treatment arms in this design can be different dose levels or different combinations of compounds.

1.1.4 Add-Arm Design for Unimodal Response

Inspired by the features of drop-the-loser design, an effective 3-stage dose-finding design that allows adding arms at the interim (“Add-arm design for unimodal response”) was proposed (Chang and Wang, 2012). This design shows benefits when the dose-response curve is unimodal.

The design starts with two arms; depending on the response of the two arms at the interim analysis and the assumption of unimodality, a decision on which new arm(s) to be added for the next stage is made. The added arm(s) will be compared with the winner arm from the first stage to decide which arm will be further carried to the end of the study. In this design, to find the best response treatment, not every treatment has to be explored; the key idea is that some inferior arms do not need to be exposed at all.
Type I error rate control and power were studied for the add-arm design. It’s shown that on average, the add-arm design is more powerful than the corresponding 2-stage drop-the-loser design. It saves sample size from drop-the-loser design by as much as 20%.

1.1.5 Other Designs

Other conventional adaptive clinical trial designs include adaptive randomization design (ARD), adaptive dose-finding design, adaptive treatment-switching design, and Bayesian adaptive design.

An adaptive randomization design is a design that allows modification of randomization schedules during the trial. One type of ARD is called response-adaptive randomization (RAR), in which the allocation probability is based on the response of the previous patients. Other adaptive randomization designs include treatment-adaptive randomization, and utility-adaptive randomization.

An adaptive dose-finding (or “dose-escalation”) design is a design at which the dose level used to treat the next-entered patient is dependent on the toxicity of the previous patients, based on traditional escalation rules. More advanced dose-escalation rules have also been proposed recently, using modeling approaches (frequentist or Bayesian framework) such as the continual reassessment method (CRM) (O’Quigley, et al. 1990). The adaptive dose-finding design can reduce the sample size and overall toxicity in a trial and improve the accuracy and precision of the estimation of maximum tolerated dose (MTD).

An adaptive treatment-switching design (ATSD) is a design that allows the investigator to switch a patient’s treatment from its initial assignment if there is evidence of lack of
efficacy or a safety concern. Despite allowing an alteration 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.

Bayesian approaches can be used in designing clinical trials for each phase. These adaptive designs are referred to “Bayesian Adaptive Design”. For example, a well-known Bayesian model-based phase I approach is continual reassessment method (CRM). Thall and Simon (1994) and Thall, Simon, and Estey (1995) introduced a class of phase II Bayesian clinical trial designs that include stopping rules based on decision boundaries for clinically meaningful events. Bayesian methods such as adaptive sample size using posterior probabilities and futility analyses using predictive probabilities can be used in confirmatory trials.

1.2 Biomarker Informed Adaptive Clinical Trial Designs

Clinical trials using conventional adaptive designs may increase the probability of success, reduce the cost and time for the drug candidate to market, and deliver the right drug to the right patient. However, the benefits of such designs might be limited if collecting measurements of primary clinical endpoint at interim takes too long. Take the oncology trials as an example: it usually takes 12 to 24 months to observe overall survival – the most commonly used and preferred regulatory primary endpoint. The extensive time needed to reach the interim analyses can present potential operational challenges (Gallo, 2006) and may delay bringing a drug to the market.
Recently, considerable interest has been drawn towards the biomarker informed adaptive clinical trial designs. A biomarker, or biological marker, refers to a measured characteristic which may be used as an indicator of the clinical measurements of disease process. That is, a short-term endpoint that is indicative of the primary clinical endpoint. Take the oncology trials as an example, the possible interim biomarker for overall survival (OS) could be: tumor size change (TSC), progression free survival (PFS), or overall response rate (ORR).

The biomarker informed adaptive clinical trial designs incorporate biomarker information at the interim stages of the study for the decisions for interim adaptation. The decisions can be made based upon the biomarker only or on the available joint information of the biomarker and the primary clinical endpoint. Friede, et al.(2011), and Liu and Pledge (2005) discussed cases where only biomarker information is used for the interim decisions. Todd and Stallard (2005), Stallard (2010) studied situations when information for both biomarker and primary endpoint is used in interim analyses.

In this dissertation, we will mainly focus on two types of biomarker informed adaptive clinical trial designs: the biomarker informed two-stage winner design and the biomarker informed add-arm design for unimodal response. Our discussions include the study design, statistical modeling, parameter estimation, error control, power evaluation and concordance analysis of the biomarker informed adaptive clinical trials.
1.2.1 Biomarker Informed Two-Stage Winner Design

Biomarker informed two-stage winner design was proposed and studied in Shun et al. (2008). Figure 1.3 shows a sketch of this design. It starts with several active treatment arms and a control arm. At interim, the inferior arms will be terminated based upon results for biomarker, and only the most promising treatment (“winner”) will be retained and carried to the end of the study with the control arm. The final comparison between the winner arm and the control arm will be performed on data from both stages and on study primary endpoint. The difference between this design and the conventional two-stage winner design is that the interim decision is based upon the biomarker.

Shun et al. (2008) studied the biomarker informed two-stage winner design with two active treatment arms and a control arm. They proposed a normal approximation for the distribution of the final test statistic of the design in order to preserve the type I error rate. However, their procedure cannot be extended to the design with more active treatment arms. In this dissertation, we will extend their work, and propose a novel statistical approach for type I error control of the biomarker informed two-stage winner design. Our method works for the design with any number of treatment arms.
1.2.2 Biomarker Informed Add-Arm Design for Unimodal Response

A new biomarker informed adaptive clinical trial design - “biomarker informed add-arm design for unimodal response” - will be proposed in this dissertation. The new design extends the idea of the add-arm design we proposed earlier (Chang and Wang, 2012). It utilizes the biomarker for early decisions on interim adaptation and the existing knowledge about the dose-response relationship to optimize the procedure of selecting best candidate treatment. The key element of this design is that, some inferior treatments could be identified, and do not need to be explored. This design is shown to be more efficient than a corresponding biomarker informed two-stage winner design via simulations.
1.3 Structure of the Dissertation

This dissertation develops novel statistical methods for biomarker informed adaptive clinical trial designs.

In Chapter 2, we propose a novel statistical approach for type I error control of the biomarker informed two-stage winner design. We use the idea of adjusting critical rejection values of the final test statistic of the design for preserving the type I error rate. The exact distribution of the final test statistic is derived and R functions for calculating the adjusted critical rejection values from the skewed distribution are developed. Our proposed method can work for the design with any number of treatment arms, which overcomes the limitation of the normal approximation method proposed by Shun et al. (2008). The critical rejection values associated with one-sided type I error rate 0.025 (w_{0.025}) for biomarker informed two-stage winner design with up to 7 active treatment groups are tabulated for easy reference.

In Chapter 3, we develop a new two-level correlation approach to model the two endpoints in a biomarker informed adaptive design. A new variable that describes the mean level correlation is incorporated in the model. Compared with the conventional model which only considers the individual level correlation, the new model could reflect the uncertainty of the historical knowledge more accurately. We illustrate the new model in the context of the biomarker informed two-stage winner design. The type I error control, power performance and parameter estimation of the design using the new model are discussed mathematically and via simulations. Our proposed two-level correlation model is shown to perform better than the conventional model.
In Chapter 4, we propose a new biomarker informed adaptive design called “biomarker informed add-arm design for unimodal response”. This design contributes to optimizing the procedure of dose-finding when a biomarker of the primary clinical endpoint exists and previous evidence indicates a unimodal dose-response relationship. One of the major advantages of the proposed design is that some inferior treatments could be identified and do not need to be explored. Consequently, the proposed design may further reduce the expected sample size while gaining information about the effective dose more quickly. We study the design using the two-level correlation model proposed in Chapter 3 and carry out extensive simulation studies to compare this design against the corresponding biomarker informed two-stage winner design in a variety of settings. The new design is shown to improve the efficiency of a clinical trial.

The discussions in Chapter 2, 3, 4 are within a theoretical framework built on the distributions of the biomarker and primary clinical endpoint without specifying an exact regression model linking the two endpoints. In Chapter 5, we further study a realistic case in non-small cell lung cancer when a model linking the two endpoints (tumor size change and overall survival) has been established. We investigate the factors that influence the concordance of decisions made based on tumor size change and overall survival. This study helps us better understand the profit and loss of using a biomarker for interim decisions.

We provide concluding remarks and discuss the possibility of using a biomarker informed adaptive design in a late phase confirmatory trial in Chapter 6.
CHAPTER 2

FINDING CRITICAL VALUES WITH CONTROLLED TYPE I ERROR FOR A BIOMARKER INFORMED TWO-STAGE WINNER DESIGN

2.1 Motivation

As sponsors recognize its value and regulators come on board, adaptive clinical trial designs are growing in popularity in recent years. Adaptive trial designs use accumulating data to allow sponsors to adjust parameters for ongoing clinical trials, including dosage, subject population or sample size. Based on adaptations applied, adaptive designs can be classified into three categories: prospective, concurrent and retrospective adaptive designs (Chow and Chang, 2008). Prospective adaptations (also known as design adaptations) imply that need for such adaptations are envisioned and approved in the protocol at the beginning of the trial. On the contrary, concurrent adaptations (also known as \textit{ad-hoc} adaptation) imply changes that could not be envisioned at the beginning, but their need became apparent as trial continues. Retrospective adaptations generally imply changes in statistical analysis plan made prior to database lock or un-blinding. The use of adaptive designs can be seen to accelerate clinical development and improve efficiency.

General concerns with using adaptive designs in drug development as listed in FDA's draft guidance on adaptive design in clinical trials are as follows:

A. Potential to increase the chance of erroneous positive conclusions and of positive study results that are difficult to interpret

1. Bias associated with the multiplicity of options
2. Difficulty in interpreting results when a treatment effect is shown

3. Operational bias

B. Potential for counterproductive impacts of adaptive design
   1. Potential to limit identifying gaps in knowledge
   2. Elimination of time to thoughtfully explore study results

C. Complex adaptive designs – potential for increased planning and more advanced time frame for planning

From statistical point of view, adaptations to trial and/or statistical procedures could introduce bias/variation to data collection, result in a shift in location and scale of the target patient population, and lead to inconsistency between hypotheses to be tested and the corresponding statistical tests.

These concerns will not only have an impact on the accuracy and reliability of statistical inference drawn on the treatment effect, but also present challenges for development of appropriate statistical methodology for an unbiased and fair assessment of the treatment effect. One of these challenges is to preserve the overall type I error rate for adaptive clinical trials at pre-specified level of significance.

A variety of approaches have been proposed over the past decade for preventing type I error inflation, some of which have generated a great deal of attention, accounting for much of the recent interest in adaptive designs. These adjustment approaches include: error spending approach for classical group sequential plans (Pocock 1977, O’Brien and Fleming 1979, Lan and DeMets 1983); Combination of p-values, such as Fisher’s combination test (Bauer 1989, Bauer and Kohne, 1994), Inverse Normal Method
(Lehmacher and Wassmer 1999), sum of p-values approach (Chang 2008); conditional error function (Proschon and Hunsberger 1995, Liu and Chi 2001 Muller and Shafer 2001); fixed weighting method (Cui, Hung and Wang 1999); variance spending method (Fisher 1998, Shen and Fisher 1999); and multiple testing methodology such as closed test procedures (Bretz, et al. 2006, Holm 1979, Hochberg 1988, Hommel 1988). These methods work for conventional adaptive designs that use the same study endpoint at the interim and the final stages of the study.

Considerable interest has been drawn towards the biomarker informed adaptive clinical trial designs recently. These designs make interim decisions based upon inference on a biomarker, which is a short-term endpoint that is indicative of the primary endpoint, and hence have the potential to prevent certain operational challenges and to bring a drug to market faster. To prevent type I error inflation for biomarker informed adaptive trials, Todd and Stallard (2005) presented a method for group sequential trials for which the interim treatment selections are based upon the biomarker only. Stallard (2010) proposed a method for group sequential trials that use both the available biomarker and primary endpoint information for treatment selections. Their method controls the type I error rate in the strong sense. Friede et al. (2011) considered a biomarker informed drop-the-losers design, they brought together combination tests for adaptive designs and the closure principle for multiple testing, and achieved control of the family-wise type I error rate in the strong sense. Scala and Glimm (2011) discussed a method for the biomarker informed adaptive designs when the endpoints are time-to-event data. Jenkins et al. (2010)
presented an approach for a biomarker informed group sequential enrichment design with time-to-event endpoints.

Shun, Lan and Soo (2008) studied a biomarker informed two-stage winner design, where interim decision is made by ranking of the observed effects of biomarker, they derived the unconditional distribution of the final test statistic for the design with two active treatment arms and one control arm, and proposed its normal approximation for calculation of the critical value to preserve type I error rate. However, the proposed normal approximation procedure cannot be extended to designs with more active treatment groups. In this chapter, we propose a novel statistical approach for biomarker informed two-stage winner design that can accommodate multiple active arms and control type I error. We use the idea of adjusting critical rejection values of the final test statistic of the design for preserving the type I error rate. The exact distribution of the final test statistic is derived and R functions for calculating the adjusted critical rejection values from the skewed distribution are developed. The critical rejection values associated with one-sided type I error rate 0.025 for biomarker informed two-stage winner design with up to 7 active treatment groups are tabulated for easy reference.

What makes a biomarker informed adaptive design special is that, two endpoints are involved in the design: the biomarker and the study primary endpoint. Therefore, the model used to describe the relationship between the two endpoints plays a role for evaluation of the design. Shun et al. (2008) used the conventional approach to model the two endpoints in a biomarker informed two-stage winner design. Wang, Chang and Menon (2013) proposed a two-level correlation model to describe the relationship
between the two endpoints, which turns to be more appropriate compared with the conventional model. In this chapter, both models will be considered in our discussion for finding critical rejection values for final test statistic to preserve type I error rate.

2.2 Finding Critical Rejection Values for Final Test Statistic in Biomarker Informed Two-Stage Winner Design

2.2.1 Biomarker Informed Two-Stage Winner Design

A biomarker informed two-stage winner design is a design that combines a phase II and a phase III study. It starts with several active treatment arms and a control arm with a planned interim analysis on biomarker. At interim, the inferior arms will be terminated based on ranking of the observed effects of biomarker, and only the most promising treatment ("winner") will be carried till the end of the study with the control arm. The final comparison between the winner arm and the control arm will be performed on data from both stages and on study primary endpoint. This design has the potential to shorten the duration of the trial for drug development and avoid unnecessary waste of resources.

Let $K$ be number of treatment groups ($K - 1$ active treatment groups, and 1 control group), and $N$ be the maximum sample size for each treatment group. Assume the interim analysis is planned at the information time $I = \frac{n_1}{N}$, where $n_1$ is the interim sample size ($n_1 < N$). Two sets of measurements are obtained: $\{X_t^{(j)} | i = 1, \ldots, n_1\}$, the measurements of the biomarker at interim stage for $i$th person in $j$th treatment group; and $\{Y_t^{(j)} | i = 1, \ldots, N\}$, the measurements of the study primary endpoint at final stage for $i$th...
person in jth treatment group. j = 0, 1, ..., K − 1. j = 0 represents the control group and j = 1, ..., K − 1 the active treatment groups.

Let $X_{n_1}^{(j)} = \frac{1}{n_1} \sum X_{i}^{(j)}$ be the mean of the biomarker measurements for treatment group j at interim, and $Y_{N}^{(j)} = \frac{1}{N} \sum Y_{i}^{(j)}$ be the mean of the primary endpoint measurements for treatment group j at final. The decision rule of the winner design is that, if the interim observations $X_{n_1}^{(j)} = \max \left( X_{n_1}^{(1)}, ..., X_{n_1}^{(K-1)} \right)$, select treatment j as the effective treatment, and carry only treatment group j and the control group to the end of the study. The option that more than one treatment groups will be kept when the interim outcomes are almost the same is not considered, because either treatment group can be selected in this situation. The final assessment will be based on the study primary endpoint Y comparing the selected treatment group j and the control group.

2.2.2 The Two Models for Fitting the Two Endpoints

In this section, we briefly review two commonly used techniques for modeling the two endpoints (i.e., the biomarker and the study primary endpoint) in a biomarker informed two-stage winner design. The first is the conventional approach, which uses the individual-level correlation $\rho$ to describe the relationship between the biomarker and the primary endpoint. The other is the two-level correlation model proposed by Wang, et al. (2013). This model considers both the individual-level and mean-level correlation between the biomarker and the primary endpoint.
We assume both the biomarker and the primary endpoint are normally distributed in the design. Let $u_j^X$ be the mean of biomarker for treatment group $j$, $\sigma_X^2$ be the variance. For a fixed $j$, assume $\{X_i^{(j)} | i = 1, ..., n_1\}$ i.i.d., and $X_i^{(j)} \sim N(u_j^X, \sigma_X^2)$. Denote the standardized mean of biomarker for each treatment group by $u_j^{X*} = \frac{u_j^X}{\sigma_X}$.

Let $u_j^Y$ be the mean of study primary endpoint for treatment group $j$, $\sigma_Y^2$ be the variance. For a fixed $j$, assume $\{Y_i^{(j)} | i = 1, ..., N\}$ i.i.d., and $Y_i^{(j)} \sim N(u_j^Y, \sigma_Y^2)$. Denote the standardized mean of study primary endpoint for treatment group $j$ by $u_j^{Y*} = \frac{u_j^Y}{\sigma_Y}$.

Assume also that $\{X_i^{(j)} | i = 1, ..., n_1\}$ and $\{Y_i^{(j)} | i = 1, ..., n_1\}$ are correlated with a correlation $\rho$ for the same $j$ and $i$, that is $\rho = \text{Corr}(X_i, Y_i)$.

Since both endpoints are assumed to be normally distributed, the conventional approach uses a multivariate normal distribution with a correlation coefficient $\rho$ for modeling the relationship between the biomarker and primary endpoint:

$$
\begin{pmatrix}
X_i^{(j)} \\
Y_i^{(j)}
\end{pmatrix} \sim N
\begin{pmatrix}
u_j^X \\
u_j^Y
\end{pmatrix}
\begin{pmatrix}
\sigma_X^2 & \rho \sigma_X \sigma_Y \\
\rho \sigma_X \sigma_Y & \sigma_Y^2
\end{pmatrix}
, j = 0, 1, ..., K - 1.
$$

This approach was used for modeling the two endpoints in the study of Shun et al. (2008), Li et al. (2010), etc. It has the limitation that the means for both endpoints have to be specified while running power simulation, which sometimes might cause problems if no solid historical knowledge is available.

The two-level correlation model proposed by Wang et al. (2013) incorporated a new variable, $R_j$, into the model, which refers to the estimated mean level correlation between
the biomarker and the primary endpoint. The new variable, together with its distribution, would reflect the uncertainty about the mean-level relationship between the two endpoints due to a small sample size of historical data.

Assume \( R_j \) is normally distributed and centered at \( r_j, r_j = \frac{u_j^{X^*}}{u_j^{X^*}}, \) which is the true mean-level correlation between the two endpoints. The two-level correlation model can be written as follows:

\[
\begin{align*}
\left( \frac{x_i^{(j)}}{\sigma_X} \right) & \sim N \left( \left( \frac{u_j^{X^*}}{R_j u_j^{X^*}} \right), (1 \quad \rho) \right) \\
\left( \frac{y_i^{(j)}}{\sigma_Y} \right) & \sim N \left( \left( \frac{u_j^{X^*}}{R_j u_j^{X^*}} \right), (1 \quad \rho) \right)
\end{align*}
\]

\( R_j \sim N(r_j, \sigma_{r_j}^2), \quad j = 0, 1, ..., K - 1. \)

And the unconditional distribution of the model could be expressed as follows (see Chapter 3 for details):

\[
\begin{align*}
\left( \frac{x_i^{(j)}}{\sigma_X} \right) & \sim N \left( \left( \frac{u_j^{X^*}}{u_j^{Y^*}} \right), \left( \begin{array}{cc} 1 & \rho \sqrt{\left( u_j^{X^*} \right)^2 \sigma_{r_j}^2 + 1} \\
\rho \sqrt{\left( u_j^{X^*} \right)^2 \sigma_{r_j}^2 + 1} & \left( u_j^{X^*} \right)^2 \sigma_{r_j}^2 + 1 \end{array} \right) \right), \quad j = 0, 1, ..., K - 1.
\end{align*}
\]

This model takes account of the uncertainty about the mean-level correlation, and have been shown to be more appropriate compared with the conventional model.

Both of the two models will be considered in the next section when we derive the distribution of the final test statistic for the biomarker informed two-stage winner design.

Designs with up to 7 active treatment arms will be studied.
2.2.3 Test Statistic and its Distribution

To prevent type I error inflation of the biomarker informed two-stage winner design, we use the idea of adjusting critical rejection values of the final test statistic of the design. In this section, we derive the exact distribution of the final test statistic for the biomarker informed two-stage winner design under the conventional one-level correlation model, and the asymptotic distribution of the final test statistic under Wang, et al.’s two-level correlation model. As it is shown, this derivation approach works for biomarker informed two-stage winner designs that accommodate any number of active arms.

Consider the following hypotheses:

\[ H_0: u_1^Y = \ldots = u_{K-1}^Y = u_0^Y \]

\[ H_1: u_1^Y > u_0^Y \text{ or } \ldots \text{ or } u_{K-1}^Y > u_0^Y \]

It is reasonable to assume that \( u_1^X = \ldots = u_{K-1}^X = u_0^X \) when \( u_1^Y = \ldots = u_{K-1}^Y = u_0^Y \) and \( \rho \neq 0 \). For simplicity, assume \( \sigma_f^2 \) is known.

Let \( G_j \) be the test statistic comparing the primary endpoint of the \( j \)th treatment group and the control group.

Under the conventional model, let

\[
G_j = \sqrt{\frac{N}{2\sigma_f^2}} \left( \bar{Y}_N^{(j)} - \bar{Y}_N^{(0)} \right)
\]

It could be shown that,

\[
G_j \sim N \left( \frac{u_1^Y - u_0^Y}{\frac{2\sigma_f^2}{\sqrt{N}}}, 1 \right)
\]

and under \( H_0 \), \( G_j \sim N(0, 1) \).
Under the two-level correlation model proposed by Wang et al. (2013), let

\[ G_j = \sqrt{\frac{N}{\sigma^2 + (u_j^{(k)})^2 \sigma^2 + (u_0^{(k)})^2 \sigma^2 + 2}} \left( Y_j^{(j)} - Y_0^{(0)} \right) \]

It could be shown that \( G_j \) is asymptotically normal,

\[ G_j \sim N \left( \sqrt{\frac{N}{(u_j^{(k)})^2 \sigma^2 + (u_0^{(k)})^2 \sigma^2 + 2}} (u^*_j - u^*_0), 1 \right) \]

And under \( H_0 \), \( G_j \) is asymptotically standard normal.

The final test statistic of the biomarker informed two-stage winner design can then be expressed as:

\[ W = G_j \text{, if } X_n^{(j)} = \max \left( X_n^{(1)}, \ldots, X_n^{(K-1)} \right) \quad j = 1, \ldots, K - 1. \]

That is, conditional on the interim selection, \( W \) takes on the value of the effect from the “winner” treatment group as the final test statistic.

For the very general case under \( H_1 \), the distribution of the final test statistic \( W \) could be derived as:

\[ F_W(w) = P(W < w) = \sum_{j=1}^{K-1} P \left( G_j < w, X_n^{(j)} = \max \left( X_n^{(1)}, \ldots, X_n^{(K-1)} \right) \right) \]

\[ = \sum_{j=1}^{K-1} P \left( G_j < w, X_n^{(j)} - X_n^{(m_1)} > 0, \ldots, X_n^{(j)} - X_n^{(m_{K-2})} > 0 \right) \]

\[ = \sum_{j=1}^{K-1} P \left( P_{j0} < a_{j0}, P_{j1} > a_{j1}, \ldots, P_{j,K-2} > a_{j,K-2} \right) \]

Under the conventional model,

\[ P_{j0} = G_j - \sqrt{\frac{N}{2 \sigma^2}} \left( u_j^* - u_0^* \right) \]
\[
P_{j1} = \frac{x_{n_1}^{(j)}}{t} - \frac{x_{n_1}^{(m_1)}}{t} - \left( \frac{u_j^X}{t} - \frac{u_{m_1}^X}{t} \right)
\]

\[
P_{j,K-2} = \frac{x_{n_1}^{(j)}}{t} - \frac{x_{n_1}^{(m_{K-2})}}{t} - \left( \frac{u_j^X}{t} - \frac{u_{m_{K-2}}^X}{t} \right)
\]

\[
a_{j0} = w - \frac{\sqrt{N}}{2\sigma^2_y}(u_j^Y - u_0^Y)
\]

\[
a_{j1} = \left( \frac{u_j^X}{t} - \frac{u_{m_1}^X}{t} \right)
\]

\[
a_{j,K-2} = -\left( \frac{u_j^X}{t} - \frac{u_{m_{K-2}}^X}{t} \right)
\]

\[
t = \frac{\sigma^2_X}{n_1}
\]

\(j \neq m_1 \neq \cdots \neq m_{K-2}, j, m_1, \ldots, m_{K-2} \in \{1, 2, \ldots, K-1\}\).

\((P_{j0}, P_{j1}, \ldots, P_{j,K-2})'\) is from a multivariate normal distribution.

If \(K = 4\):

\[
\begin{pmatrix}
P_{j0} \\
P_{j1} \\
P_{j2} \\
P_{j3}
\end{pmatrix}
\sim N\left(\begin{pmatrix}0 \\ 0 \\ 0 \\ 0\end{pmatrix}, \gamma \begin{pmatrix}1 & \gamma & \gamma \\ \gamma & 2 & 1 \\ \gamma & 1 & 2 \end{pmatrix}\right),\text{ where } \gamma = \frac{n_1}{\sqrt{2N}}\rho.
\]

If \(K = 5\):

\[
\begin{pmatrix}
P_{j0} \\
P_{j1} \\
P_{j2} \\
P_{j3}
\end{pmatrix}
\sim N\left(\begin{pmatrix}0 \\ 0 \\ 0 \\ 0 \\
\end{pmatrix}, \gamma \begin{pmatrix}1 & \gamma & \gamma & \gamma \\ \gamma & 2 & 1 & 1 \\ \gamma & 1 & 2 & 1 \\ \gamma & 1 & 1 & 2 \end{pmatrix}\right).
\]
More generally, for any value of $K$ ($K \geq 3$):

\[
\begin{pmatrix}
P_{j0} \\
P_{j1} \\
P_{j2} \\
\vdots \\
P_{j,K-2}
\end{pmatrix}
\sim N
\begin{pmatrix}
0 \\
0 \\
\vdots \\
0
\end{pmatrix},
\begin{pmatrix}
1 & \gamma & \cdots & \gamma \\
\gamma & 2 & \cdots & 1 \\
\vdots & \vdots & \ddots & \vdots \\
\gamma & \cdots & \cdots & 1
\end{pmatrix}
\]

Under the two-level correlation model proposed by Wang et al. (2013),

\[
P_{j0} = G_j - \sqrt{\frac{N}{\left(\frac{\sigma_j^2}{n_1}\right)}} \left(\sqrt{n_1}u_j^* - \sqrt{n_1}u_{m_1}^*ight)
\]

\[
P_{j1} = \frac{x_{n_1}^{(j)} - x_{n_1}^{(m_1)}}{\sqrt{\frac{\sigma_j^2}{n_1}}} - \left(\sqrt{n_1}u_j^* - \sqrt{n_1}u_{m_1}^*ight)
\]

\[
P_{j,K-2} = \frac{x_{n_1}^{(j)} - x_{n_1}^{(m_{K-2})}}{\sqrt{\frac{\sigma_j^2}{n_1}}} - \left(\sqrt{n_1}u_j^* - \sqrt{n_1}u_{m_{K-2}}^*ight)
\]

\[
a_{j0} = w - \sqrt{\frac{N}{\left(\frac{\sigma_j^2}{n_1}\right)}} \left(\sqrt{n_1}u_j^* - \sqrt{n_1}u_{m_1}^*ight)
\]

\[
a_{j1} = -\left(\sqrt{n_1}u_j^* - \sqrt{n_1}u_{m_1}^*\right)
\]

\[
a_{j,K-2} = -\left(\sqrt{n_1}u_j^* - \sqrt{n_1}u_{m_{K-2}}^*\right)
\]

\(j \neq m_1 \neq \cdots \neq m_{K-2}, \ j, m_1, \ldots, m_{K-2} \in \{1, 2, \ldots, K - 1\}\).

\([P_{j0}, P_{j1}, \ldots, P_{j,K-2}]'\) is approximately from a multivariate normal distribution.
If $K = 4$:

\[
\begin{pmatrix}
    P_{j0} \\
    P_{j1} \\
    P_{j2} \\
    P_{j3}
\end{pmatrix}
\sim N
\begin{pmatrix}
    \begin{pmatrix}
        0 & \gamma_j & \gamma_j & \gamma_j \\
        \gamma_j & 2 & 1 & 1 \\
        \gamma_j & 1 & 2 & 1 \\
        \gamma_j & 1 & 1 & 2
    \end{pmatrix}
\end{pmatrix},
\text{where } \gamma_j = \rho \sqrt{\frac{\sigma_j^2}{N}} \left( \frac{(u_j^y)^2 \sigma_j^2 + 1}{(u_j^y)^2 \sigma_j^2 + (u_j^x)^2 \sigma_j^2 + \sigma_{\mu_0}^2} \right).
\]

If $K = 5$:

\[
\begin{pmatrix}
    P_{j0} \\
    P_{j1} \\
    P_{j2} \\
    P_{j3} \\
    P_{j4}
\end{pmatrix}
\sim N
\begin{pmatrix}
    \begin{pmatrix}
        0 & \gamma_j & \gamma_j & \gamma_j & \gamma_j \\
        \gamma_j & 2 & 1 & 1 & 1 \\
        \gamma_j & 1 & 2 & 1 & \vdots \\
        \gamma_j & 1 & \vdots & \vdots & 1 \\
        \gamma_j & 1 & \ldots & 1 & 2
    \end{pmatrix}
\end{pmatrix}.
\]

More generally, for any value of $K \geq 3$:

\[
\begin{pmatrix}
    P_{j0} \\
    P_{j1} \\
    P_{j2} \\
    \vdots \\
    P_{j,K-2}
\end{pmatrix}
\sim N
\begin{pmatrix}
    \begin{pmatrix}
        0 & \gamma_j & \gamma_j & \ldots & \gamma_j \\
        \gamma_j & 2 & 1 & \ldots & 1 \\
        \gamma_j & 1 & 2 & \ddots & \vdots \\
        \gamma_j & \ddots & \ddots & \ddots & \vdots \\
        \gamma_j & 1 & \ldots & 1 & 2
    \end{pmatrix}
\end{pmatrix}.
\]

Thus, for a biomarker informed two-stage winner design with $K \geq 3$ treatments, the distribution of final test statistic under the conventional model is:

\[
F_w(w) = \sum_{j=1}^{K-1} \int_{-\infty}^{w_j} \left( \frac{\sqrt{N} (u_j^y - u_0^y)}{2 \sigma_j^y} \right) \left( \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(p_{j0}, \ldots, p_{j,K-2}) dp_{j,K-2} \cdots dp_{j0} \right)
\]

where:
$j \neq m_1 \neq \cdots \neq m_{K-2}, \ j, m_1, \ldots, m_{K-2} \in \{1, 2, \ldots, K-1\};$

\[
f(p_{j0}, \ldots, p_{jK-2}) \text{ is p.d.f. of } N \begin{pmatrix}
0 \\
0 \\
\vdots \\
0
\end{pmatrix}
, \begin{pmatrix}
1 & \gamma & \gamma & \cdots & \gamma \\
\gamma & 2 & 1 & \cdots & 1 \\
\gamma & 1 & 2 & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
\gamma & 1 & \cdots & 1 & 2
\end{pmatrix}
, \text{ and } \gamma = \frac{n_1}{\sqrt{2N}} \rho.
\]

And the distribution of final test statistic under the two-level correlation model is:

\[
F_w(w) = \sum_{j=1}^{K-1} \int_{-\infty}^{w} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f_j(p_{j0}, \ldots, p_{jK-2}) dp_{j0} \ldots dp_{jK-2}
\]

where:

\[
j \neq m_1 \neq \cdots \neq m_{K-2}, \ j, m_1, \ldots, m_{K-2} \in \{1, 2, \ldots, K-1\};
\]

\[
f_j(p_{j0}, \ldots, p_{jK-2}) \text{ is p.d.f. of } N \begin{pmatrix}
0 \\
0 \\
\vdots \\
0
\end{pmatrix}
, \begin{pmatrix}
1 & \gamma_j & \gamma_j & \cdots & \gamma_j \\
\gamma_j & 2 & 1 & \cdots & 1 \\
\gamma_j & 1 & 2 & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
\gamma_j & 1 & \cdots & 1 & 2
\end{pmatrix}
\]

\[
\gamma_j = \rho \sqrt{\frac{n_1}{N}} \sqrt{\frac{(u_{j0}^T)^2 \sigma_{\gamma}^2 + 1}{(u_{j0}^T)^2 \sigma_{\gamma}^2 + 2}}.
\]

It can be seen that, the interim treatment selection of the design skewed the distribution of its final test statistic. Hence, appropriate statistical adjustment is necessary in order to preserve the type I error rate of the design. As the general distribution of the final test statistic is written, the type I error rate of the design can be preserved by adjusting the
critical rejection value for the final test statistic. In the following section, we focus our work on finding the adjusted critical rejection values for final test statistic.

2.2.4 Critical Rejection Values and \( R \)

Under \( H_0 \), the distribution of the final test statistic of a biomarker informed two-stage winner design with \( K \) treatments can be written as follows.

Under conventional model:

\[
F_W(w) = \sum_{j=1}^{K-1} P(P_{j0} < w, P_{j1} > 0, ..., P_{j,K-2} > 0) = (K - 1) \int_{-\infty}^{w} \int_{0}^{+\infty} \int_{0}^{+\infty} f(p_{j0}, p_{j1}, ..., p_{j,K-2}) dp_{j0} ... dp_{j,K-2}
\]

where \( f(p_{j0}, p_{j1}, ..., p_{j,K-2}) \) is p.d.f. of

\[
N \left( \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \gamma & \gamma & \cdots & \gamma \\ \gamma & 2 & 1 & \cdots & 1 \\ \gamma & 1 & 2 & \cdots & \gamma \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma & 1 & \cdots & 1 & 2 \end{pmatrix}_{(K-1)\times(K-1)} \right)
\]

and \( \gamma = \sqrt{\frac{n_1}{2N}} \rho \).

Under the two-level correlation model proposed by Wang et al. (2013):

\[
F_W(w) = \sum_{j=1}^{K-1} P(P_{j0} < w, P_{j1} > 0, ..., P_{j,K-2} > 0) = \sum_{j=1}^{K-1} \int_{-\infty}^{w} \int_{0}^{+\infty} \int_{0}^{+\infty} f_j(p_{j0}, p_{j1}, ..., p_{j,K-2}) dp_{j0} ... dp_{j,K-2}
\]

where \( f_j(p_{j0}, p_{j1}, ..., p_{j,K-2}) \) is p.d.f. of
If \( \sigma_{r_0}^2 = \sigma_{r_1}^2 = \cdots = \sigma_{r,K-1}^2 \), that is the variability of the estimated mean-level correlation for each treatment group is the same, the above distribution can be approximated by:

\[
F_W(w) = \sum_{j=1}^{K-1} P(P_{j_0} < w, P_{j_1} > 0, ..., P_{j,K-2} > 0)
\]

\[
= (K - 1) \int_{-\infty}^{w} \int_{0}^{+\infty} \cdots \int_{0}^{+\infty} f(p_{j_0}, p_{j_1}, ..., p_{j,K-2}) dp_{j_0} \cdots dp_{j,K-2}
\]

where \( f(p_{j_0}, p_{j_1}, ..., p_{j,K-2}) \) is p.d.f. of

\[
N \left( \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \end{pmatrix}_{(K-1)\times 1} , \begin{pmatrix} 1 & y_j & y_j & \cdots & y_j \\ y_j & 2 & 1 & \cdots & 1 \\ y_j & 1 & 2 & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 1 \\ y_j & 1 & \cdots & 1 & 2 \\ \end{pmatrix}_{(K-1)\times(K-1)} \right)
\]

and \( y = \sqrt{\frac{n_1}{2N}} \rho \).

Denote the distribution of final test statistic \( W \) under \( H_0 \) by \( F_0 \). Let \( w_\alpha \) be the upper 100\( \alpha \) percent quintile of \( F_0 \),

\[
w_\alpha = F_W^{-1}(1 - \alpha | H_0) = F_0^{-1}(1 - \alpha)
\]

The type I error rate of the design can be controlled at level \( \alpha \) if the 1-sided rejection region is \( \Omega = \{W: W > w_\alpha\} \).
We develop R functions for calculating the critical rejection values \( w_\alpha \) for biomarker informed two-stage winner design with up to 7 active treatment arms (Appendix A).

R functions `convention_cv_Kk (k = 4, ..., 8)` are for calculating the integration of \( F_0(w) \) for biomarker informed two-stage winner design with \( k - 1 \) active treatment groups and a control group under conventional model and therefore can be used to find critical rejection values for the final test statistic. For example, to find the critical rejection value that preserves type I error rate at 0.025 for a biomarker informed two-stage winner design with 3 active treatment arms and a control arm, `convention_cv_K4` can be used, and an output of 0.975 is expected for the critical rejection value. If the interim sample size of the design is 50, the maximum sample size per treatment group is 100, and \( \rho = 0.8 \),

```R
> convention_cv_K4(w=2.232, n1=50, N=100, rho=0.8, alpha=0.025)
```

```
$ p 
[1] 0.9750554
```

```
attr("error")
[1] 8.166802e-07
```

```
attr("status")
[1] "normal completion"
```

```
$error
[1] 5.54042e-05
```

Thus, the critical rejection value \( w_\alpha = 2.232 \) in this case.

In some scenarios, the individual-level correlation \( \rho \) is unknown, the sample correlation coefficient \( \hat{\rho} \) is suggested to be used for an approximate value of \( w_\alpha \). Simulations are
further suggested to be performed to ensure that the type I error rate of the design is preserved.

R functions \texttt{wang\_cv\_Kk} \((k = 4, \ldots, 8)\) are for calculating values for biomarker informed two-stage winner design with \(k - 1\) active treatment groups and a control group under the two-level correlation model proposed by Wang et al (2013). These functions work in a similar way as functions \texttt{convention\_cv\_Kk}. If there are unknown parameters incorporated in the functions, the parameter estimates are suggested to be used for calculating an approximate critical rejection value. Further simulations are encouraged to ensure a preserved type I error rate.

Table 2.1 – Table 2.5 provide the critical rejection values \(w_{0.025}\) for biomarker informed two-stage winner design with up to 7 active treatment group under conventional model. As expected, the critical rejection value \(w_{0.025}\) increases as \(\rho\) increases. It can also be seen that, the later the interim look, the larger the critical rejection value \(w_{0.025}\) need to be. Besides, the more active treatment groups included in the design, the larger the critical rejection value \(w_{0.025}\) have to be.
Table 2.1 \( w_{0.025} \) for biomarker informed two-stage winner design \((K = 4)\)

<table>
<thead>
<tr>
<th>( \rho )</th>
<th>0</th>
<th>0.2</th>
<th>0.5</th>
<th>0.8</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n_i/N = 0.3 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( w_{0.025} )</td>
<td>1.96</td>
<td>2.023</td>
<td>2.108</td>
<td>2.182</td>
<td>2.225</td>
</tr>
<tr>
<td>( n_i/N = 0.5 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( w_{0.025} )</td>
<td>1.96</td>
<td>2.041</td>
<td>2.146</td>
<td>2.232</td>
<td>2.279</td>
</tr>
<tr>
<td>( n_i/N = 0.8 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( w_{0.025} )</td>
<td>1.96</td>
<td>2.061</td>
<td>2.186</td>
<td>2.281</td>
<td>2.328</td>
</tr>
</tbody>
</table>

Table 2.2 \( w_{0.025} \) for biomarker informed two-stage winner design \((K = 5)\)

<table>
<thead>
<tr>
<th>( \rho )</th>
<th>0</th>
<th>0.2</th>
<th>0.5</th>
<th>0.8</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n_i/N = 0.3 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( w_{0.025} )</td>
<td>1.96</td>
<td>2.038</td>
<td>2.142</td>
<td>2.232</td>
<td>2.286</td>
</tr>
<tr>
<td>( n_i/N = 0.5 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( w_{0.025} )</td>
<td>1.96</td>
<td>2.058</td>
<td>2.188</td>
<td>2.294</td>
<td>2.352</td>
</tr>
<tr>
<td>( n_i/N = 0.8 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( w_{0.025} )</td>
<td>1.96</td>
<td>2.083</td>
<td>2.237</td>
<td>2.356</td>
<td>2.415</td>
</tr>
</tbody>
</table>

Table 2.3 \( w_{0.025} \) for biomarker informed two-stage winner design \((K = 6)\)

<table>
<thead>
<tr>
<th>( \rho )</th>
<th>0</th>
<th>0.2</th>
<th>0.5</th>
<th>0.8</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n_i/N = 0.3 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( w_{0.025} )</td>
<td>1.96</td>
<td>2.048</td>
<td>2.165</td>
<td>2.27</td>
<td>2.33</td>
</tr>
<tr>
<td>( n_i/N = 0.5 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( w_{0.025} )</td>
<td>1.96</td>
<td>2.072</td>
<td>2.218</td>
<td>2.34</td>
<td>2.408</td>
</tr>
<tr>
<td>( n_i/N = 0.8 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( w_{0.025} )</td>
<td>1.96</td>
<td>2.1</td>
<td>2.275</td>
<td>2.411</td>
<td>2.48</td>
</tr>
</tbody>
</table>
Table 2.4 $w_{0.025}$ for biomarker informed two-stage winner design ($K = 7$)

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>0</th>
<th>0.2</th>
<th>0.5</th>
<th>0.8</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n_i/N = 0.3$</td>
<td>$w_{0.025}$</td>
<td>1.96</td>
<td>2.055</td>
<td>2.185</td>
<td>2.299</td>
</tr>
<tr>
<td>$n_i/N = 0.5$</td>
<td>$w_{0.025}$</td>
<td>1.96</td>
<td>2.082</td>
<td>2.242</td>
<td>2.376</td>
</tr>
<tr>
<td>$n_i/N = 0.8$</td>
<td>$w_{0.025}$</td>
<td>1.96</td>
<td>2.112</td>
<td>2.305</td>
<td>2.455</td>
</tr>
</tbody>
</table>

Table 2.5 $w_{0.025}$ for biomarker informed two-stage winner design ($K = 8$)

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>0</th>
<th>0.2</th>
<th>0.5</th>
<th>0.8</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n_i/N = 0.3$</td>
<td>$w_{0.025}$</td>
<td>1.96</td>
<td>2.063</td>
<td>2.202</td>
<td>2.323</td>
</tr>
<tr>
<td>$n_i/N = 0.5$</td>
<td>$w_{0.025}$</td>
<td>1.96</td>
<td>2.091</td>
<td>2.262</td>
<td>2.407</td>
</tr>
<tr>
<td>$n_i/N = 0.8$</td>
<td>$w_{0.025}$</td>
<td>1.96</td>
<td>2.123</td>
<td>2.33</td>
<td>2.492</td>
</tr>
</tbody>
</table>
2.3 Discussion

In this chapter, we have proposed a novel statistical approach for type I error control of the biomarker informed two-stage winner design. We use the idea of adjusting critical rejection values of the final test statistic of the design for preserving the type I error rate. We derive the exact distribution of the final test statistic for the biomarker informed two-stage winner design under the conventional one-level correlation model, and the asymptotic distribution of the final test statistic under Wang et al.’s two-level correlation model, hence the critical rejection values \( w_\alpha \) could be obtained through mathematical integration. We developed R functions for calculating the adjusted critical rejection values from the skewed distribution of final test statistic. As shown, the critical rejection value \( w_{0.025} \) increases as any of \( \rho, k, \frac{n_1}{N} \) increases.

Our proposed method overcomes the limitation of the normal approximation method proposed by Shun et al. (2008), and works for designs with any number of treatment arms. However, it has the limitation that it works only for the biomarker informed two-stage winner design with normal interim and final endpoints. For the designs with non-normal endpoints, transformations might be used to convert the data to follow normal distribution. Developing novel approaches for type I error control for biomarker informed two-stage winner design with non-normal endpoints would be an interesting topic for future studies.
CHAPTER 3
MODELING THE BIOMARKER INFORMED ADAPTIVE
CLINICAL TRIAL DESIGNS

3.1 Introduction

The pharmaceutical industry has increased its commitment to personalized medicine as
the cost of genetic sequencing declines. Personalized medicine aims to deliver the right
drug to the right patient. In addition to discovering the right drug, finding the right dose
and identifying the right patient, it is also desirable to shorten the time of drug
development in order to bring the drug to the patient faster. Different clinical trial designs
have been proposed for this purpose. One such design is a seamless phase II/III drop-the-
losers (or “pick-the-winner”) design, which has the potential to terminate the inferior
treatment groups (i.e. the “losers”) early if no efficacy is shown. It minimizes “white
space” between phase II and phase III of the studies, and efficiently uses all the patient
data both in the learning and confirming phases.

Statistical methods exist for controlling the type I error rate and constructing estimators
for the drop-the-losers design. Stallard and Todd (2003), Kelly et. al. (2005), and Stallard
considered the method of combining p-values from different stages; Sampson and Sill
(2005) developed a uniformly most powerful conditionally unbiased test (UMPCU) for
normally distributed data; Chang, Chow and Pong (2006) suggested the contrast test with
a p-value combination method.
In drop-the-losers clinical trials, the same endpoint is used for both the interim and final analyses of the study. However, the benefits of such an approach could be limited if collecting measurements of primary clinical endpoint at interim takes very long. For example, in oncology trials, it usually takes 12 to 24 months to observe overall survival – the most commonly used and preferred regulatory primary endpoint. The extensive time needed to reach the interim analyses can present potential operational challenges (Gallo, 2006) and may delay bringing a drug to the market.

Considerable interest has been drawn towards the short-term endpoint (“biomarker”) informed adaptive seamless phase II/III designs. These designs incorporate biomarker information at the interim stages of the study. The decision(s) on interim adaptation can be made based upon the biomarker only or on the available joint information of the biomarker and the primary endpoint.

Todd and Stallard (2005) presented a biomarker informed group sequential design for which the interim treatment selection is based upon a biomarker. Stallard (2010) studied a biomarker informed group sequential design that uses both the available biomarker and primary endpoint information for treatment selection. He proposed a method for the adjustment of the usual group sequential boundaries to maintain strong control of the family-wise type I error rate. Shun, Lan and Soo (2008) presented a biomarker informed two-stage winner design with two active treatments arms and a control arm. They derived the unconditional distribution of the final test statistic for the design and proposed a normal approximation approach for the final distribution to preserve the type I error rate. Liu and Pledger (2005), Li and Wang and Ouyang (2009), and Li, et al. (2009)
considered biomarker informed drop-the-losers designs where more than one treatment can be selected at the interim. Friede et al. (2011) proposed a type I error control procedure for biomarker informed drop-the-losers design. Their method brought together combination tests for adaptive designs and the closure principle for multiple testing, which combined p-values from hypothesis tests for primary endpoint. However, the correlation between the biomarker and primary endpoint was not considered in their approach while investigating type I error rate of the design. Scala and Glimm (2011) discussed application of the biomarker informed adaptive designs when the endpoints are time-to-event data. Jenkins et al. (2010) proposed a biomarker informed group sequential enrichment design with time-to-event endpoints, and methodology was presented which controls the type I error rate.

Biomarker informed adaptive clinical trial designs have the potential to speed up drug discovery and shorten time to market. It could be very helpful for the development of personalized medicine, if the biomarker used at the interim is a good indicator of the primary clinical endpoint.

### 3.2 Motivations and Concepts

To conduct a drug trial that uses biomarker informed adaptive procedures, statistical simulations are suggested to be performed first in order to understand the operating characteristics, including sample size for a target power, of the design. An important component of this process is to specify a model for simulation of the two endpoints incorporated in the study. The conventional approach uses the one-level (‘individual-
level” only) correlation model, together with historical knowledge, to describe the relationship between the biomarker and primary endpoint. This approach can easily misestimate the power of a biomarker informed design if there is no well-established knowledge about how the biomarker and the primary endpoint are correlated. For example, when the rank order of mean responses of the biomarker for each treatment group is assumed to be the same as that of the primary endpoint based on historical observations, the power of a biomarker informed two-stage winner design is very possible to be overestimated by the conventional model, as the uncertainty of the historical knowledge has been ignored. In this case, the sample size suggested by simulation may lead to an underpowered trial.

The approval rate for new drug applications (NDAs) submitted to the FDA recently is only about 40%. This fact indicates that there are plenty of trials that are underpowered. It is desirable to propose approaches that lead to a more accurate assessment of clinical trial designs.

In this manuscript, we propose a two-level correlation model to fit the two endpoints in a biomarker informed adaptive design. This model considers not only the individual-level correlation between the biomarker and the primary endpoint, but also accounts for the variability of the estimated mean-level correlation (or “mean-level association”). The uncertainty due to a small sample size of historical data about the relationship between the two endpoints could be more accurately reflected in our model.

The new model is illustrated in the context of a two-stage winner design with three active treatment arms and a control arm. We assume both the biomarker and primary endpoint
are normally distributed, and derive the asymptotic distribution of the final test statistic of the design using our new model. We further propose a statistical approach for type I error control of the design, and carry out extensive simulations to study the power performance of the design for different relationship between the biomarker and primary endpoint. It is shown that the new proposed model performs better than the conventional one-level correlation model. An approach for estimating the parameters incorporated in our new model is further proposed.

Throughout this chapter, we want to deliver the message that the conventional one-level correlation model is not sufficient for modeling the relationship between the two endpoints in a biomarker informed adaptive design. The shape of mean-level correlation, as well as the uncertainty about the shape, need also be considered.

### 3.3 Issues in Conventional One-Level Correlation Model

The conventional one-level correlation model used for describing the relationship of the two endpoints in a biomarker informed adaptive design considers only the individual-level correlation ($\rho$).

For example in a biomarker informed adaptive clinical trial design, let $\{X_i \mid i = 1, \ldots, n_1\}$ be measurements of interim endpoint, and $\{Y_i \mid i = 1, \ldots, N\}$ be measurements of primary endpoint, assume both endpoints are normally distributed, then the conventional approach will use the following multivariate normal distribution with a correlation coefficient $\rho$, to fit the two endpoints:
\[(X_i, Y_i) \sim N \left( \begin{pmatrix} u_X^i \\ u_Y^i \end{pmatrix}, \begin{pmatrix} \sigma_X^2 & \rho \sigma_X \sigma_Y \\ \rho \sigma_X \sigma_Y & \sigma_Y^2 \end{pmatrix} \right). \]

Hence if the conventional one-level correlation model is used in statistical simulations for biomarker informed trials, the means of the two endpoints have to be specified based on the historical knowledge (as shown in Li et al. 2009). In this way, if there is no well-established historical knowledge about the two endpoints, the simulation study results might be inaccurate. For example, if only a small amount of historical data is available for the two endpoints in a biomarker informed two-stage winner trial, which indicates a same rank order of the mean responses of the two endpoints for each treatment group, then the power of this trial is very possible to be overestimated by the conventional model, as the uncertainty of the historical knowledge has been ignored. In addition, there would not be much difference in power between different values of correlation coefficient \(\rho\) between the two endpoints.

Friede et al. (2011) pointed out that the effect of the individual level correlation \(\rho\) between the biomarker and primary endpoint on power is small if the means of biomarker in treatment groups are fixed and are different. In their paper, the authors showed (in Figure 2(a)) how the estimated power of a biomarker informed design has changed for different estimations of treatment difference in biomarker. As shown in their figure, there is almost no difference in the power of the design for different values of the correlation coefficient \(\rho\). For example, when the estimated treatment difference in biomarker is 0.2, the estimated power of their design changes from around 83% to 85% as \(\rho\) increases from 0 to 1. Li et al. (2009) also mentioned that the influence of \(\rho\) on power is really small.
when compared to other factors. Their paper showed an increase in simulated power from 70.5% to 73.7% as ρ increased from 0.2 to 0.8. We ran simulations for a two-stage winner design with a survival primary endpoint. The design was assumed to have 3 active treatment arms and 1 control arm, with fixed survival means 2.46, 3.32, 3 and 2.22 for each treatment arm respectively; and fixed biomarker log-means 1.6, 1.9, 1.8 and 1.5. The critical value for the final test statistic was obtained by simulation with the type I error rate at 0.05 level. Our simulation results (Table 3.1) are consistent with the previous findings.

Table 3.1 Power Evaluation Using Conventional One-Level Correlation Model

<table>
<thead>
<tr>
<th>ρ</th>
<th>censoring rate</th>
<th>Interim size</th>
<th>Max size</th>
<th>power</th>
<th>power with best-treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.2</td>
<td>90</td>
<td>270</td>
<td>94.0%</td>
<td>72.2%</td>
</tr>
<tr>
<td>0.3</td>
<td>0.2</td>
<td>90</td>
<td>270</td>
<td>95.0%</td>
<td>72.8%</td>
</tr>
<tr>
<td>0.5</td>
<td>0.2</td>
<td>90</td>
<td>270</td>
<td>95.8%</td>
<td>73.8%</td>
</tr>
<tr>
<td>0.8</td>
<td>0.2</td>
<td>90</td>
<td>270</td>
<td>96.0%</td>
<td>73.6%</td>
</tr>
</tbody>
</table>

It could be seen from Table 3.1 that if the conventional one-level correlation approach is used, when ρ increase from 0 to 0.8, the simulated power of the two-stage winner design changes from 94.0% to 96.0%. It is only slightly different. This finding violates the presumption that the design should have a better performance when ρ is large, as ρ is the only index in the conventional model to describe the correlation between the two
endpoints. In addition, the simulation results also suggest that the power is probably overestimated by the conventional one-level correlation model.

Figure 3.1 illustrates different cases where the individual level correlation $\rho$ between the biomarker and primary endpoint is the same. Two different biomarkers are described in the figure, both of which are correlated with the primary endpoint with correlation coefficient $\rho = 0.9$. Consider designs that use the two biomarkers at interim respectively: if the same treatment difference is found on biomarkers at interim, the power of the two designs might be different, since the treatment difference on the primary endpoint is different due to the different slope. Therefore, it’s not sufficient to describe the relationship between biomarker and primary endpoint by only considering the individual level correlation $\rho$. The slope, which is the rate of change of the primary endpoint with respect to biomarker, should also be incorporated. The conventional one-level correlation model does not incorporate the variability of the slope caused by the uncertainty of historical data, which might easily lead to misestimated power.
3.4 Biomarker Informed Two-Stage Winner Design with the Proposed Two-Level Correlation Model

We propose a two-level correlation model to describe the relationship between the two endpoints in a biomarker informed adaptive design, in order to obtain a more accurate assessment for the design. To illustrate the proposed model, we consider a biomarker informed two-stage winner design with 3 active treatment arms and a control arm.

3.4.1 The Biomarker Informed Two-Stage Winner Design

A biomarker informed two-stage winner design is a special biomarker informed drop-the-losers design, where only a single active treatment will be selected for final evaluation. It combines a phase II and a phase III study, and starts with several active treatment arms and a control arm with a planned interim analysis. At the interim, the inferior treatments will be dropped based on observations of the biomarker, which is a short-term endpoint that is indicative of the primary endpoint. Only the most promising one will be retained and carried until the end of the study. The final comparison of winner treatment with control is performed on the data collected from both stages. This design could potentially shorten the duration of the development, and aid in delivering the medicine to the patient faster.

For a biomarker informed two-stage winner design, Let \( \{X_{i}^{(j)} | i = 1, \ldots, n_1\} \) be measurements of biomarker at interim stage, and \( \{Y_{i}^{(j)} | i = 1, \ldots, N\} \) be the measurements of primary endpoint obtained at final stage. \( n_1 \) is the interim sample size per group, and
$N$ is the maximum sample size for each treatment, $n_1 < N$; $j = 0, 1, 2, 3$. Let $j = 0$ represents the control group and $j = 1, 2, 3$ the 3 active treatment groups. $\overline{X}_{n_1}^{(j)} = \frac{1}{n_1} \sum X_i^{(j)}$ is the mean of the biomarker measurements for treatment group $j$, and $\overline{Y}_N^{(j)} = \frac{1}{N} \sum Y_i^{(j)}$ is the mean of the primary endpoint measurements for treatment group $j$.

The decision rule of the winner design is that, if the interim observations $\overline{X}_{n_1}^{(j)} = \max \left( \frac{X_{n_1}^{(1)}}{\sigma^2}, \frac{X_{n_1}^{(2)}}{\sigma^2}, \frac{X_{n_1}^{(3)}}{\sigma^2} \right)$, select treatment $j$ as the effective treatment, and carry only treatment group $j$ and the control group to the end of the study. The option that more than one treatment groups will be kept when the interim outcomes are almost the same is not considered, because either treatment group can be selected in this situation. The final assessment will be based on the study primary endpoint $Y$ comparing the selected treatment group $j$ and the control group.

### 3.4.2 The Proposed Two-Level Correlation Model

For simplicity, we assume both the biomarker and the primary endpoint are normally distributed in the biomarker informed two-stage winner design. The two-level correlation approach we proposed models the design in the following way:

Assume, $u_j^{X^*} \sim 0, 1, 2, 3$, are standardized means of biomarker for treatment group $j$, and $\sigma^2_X$ is the common variance. For a fixed $j$, assume $\{X_i^{(j)} | i = 1, ..., n_1 \}$ i.i.d., and $\frac{X_i^{(j)}}{\sigma_X} \sim N(u_j^{X^*}, 1)$. 

Assume \( \{ Y^{(j)}_i \mid i = 1, \ldots, N \} \) i.i.d., denote the true mean-level correlation between biomarker and primary endpoint by \( \tau_j \) for treatment group \( j \). If \( \tau_j \) is known, the conditional distribution \( \frac{Y^{(j)}_i}{\sigma_y} \mid \tau_j \sim N(u^x_j, 1) \), where \( \tau_j = \frac{u^x_j}{u^x_j} \).

In reality, since the true mean-level correlation \( \tau_j \) is unknown, an estimate \( R_j \) is obtained from historical data to describe the estimated mean-level correlation. We assume \( R_j \sim N(\tau_j, \sigma^2_{\tau_j}) \), and hence \( \frac{Y^{(j)}_i}{\sigma_y} \mid R_j \sim N(R_j u^x_j, 1) \).

It is easy to show, under this setting, the unconditional distribution for the primary endpoint is: \( \frac{Y^{(j)}_i}{\sigma_y} \sim N(u^x_j, (u^x_j)^2 \sigma^2_{\tau_j} + 1) \).

The individual level correlation between the biomarker and primary endpoint is denoted by \( \rho \), where \( \rho = \text{Corr}(X_i, Y_i) \).

Thus, the conditional joint two-level correlation model of biomarker and primary endpoint can be written as:

\[
\begin{pmatrix}
\frac{X^{(j)}_i}{\sigma_x} \\
\frac{Y^{(j)}_i}{\sigma_y} \\
\end{pmatrix}
\sim
N
\begin{pmatrix}
\begin{pmatrix}
u^x_j \\
u^y_j \\
\end{pmatrix}
\begin{pmatrix}
1 & \rho \\
R_j & u^x_j \\
\end{pmatrix}
\end{pmatrix}
\]

\( R_j \sim N(\tau_j, \sigma^2_{\tau_j}) \)

and the unconditional joint distribution could be expressed as:

\[
\begin{pmatrix}
\frac{X^{(j)}_i}{\sigma_x} \\
\frac{Y^{(j)}_i}{\sigma_y} \\
\end{pmatrix}
\sim
N
\begin{pmatrix}
\begin{pmatrix}
u^x_j \\
u^y_j \\
\end{pmatrix}
\begin{pmatrix}
1 & \rho \sqrt{(u^x_j)^2 \sigma^2_{\tau_j} + 1} \\
\rho \sqrt{(u^x_j)^2 \sigma^2_{\tau_j} + 1} & (u^x_j)^2 \sigma^2_{\tau_j} + 1 \\
\end{pmatrix}
\end{pmatrix}
\], \( j = 0, 1, 2, 3. \)
The new variable $R_j$, incorporated in the model, and its distribution accounts for the mean-level correlation between biomarker and primary endpoint. The uncertainty, due to a small sample size of historical data, about the relationship between the two endpoints should be reflected in the model in this way. The conventional one-level correlation model treats $R_j$ as the true mean-level correlation $r_j$ for simulation, which might be inappropriate.

3.4.3 Test Statistic and its Distribution

Consider the following hypotheses:

$H_0: u_1^* = u_2^* = u_3^* = u_0^*$

$H_1: u_1^* > u_0^* \text{ or } u_2^* > u_0^* \text{ or } u_3^* > u_0^*$

We want to test if there is any treatment that shows significantly better efficacy than the control group.

It is reasonable to assume that $u_1^* = u_2^* = u_3^* = u_0^*$ when $u_1^* = u_2^* = u_3^* = u_0^*$ and $\rho \neq 0$. For simplicity, assume $\sigma_\rho^2$ is known.

Let the test statistic comparing the primary endpoint of the jth treatment group and the control group be:

$$G_j = \sqrt{\frac{N}{\sigma_\rho^2 \left( u_j^* \right)^2 + \left( u_0^* \right)^2 + 2\sigma_\rho^2}} \left( \frac{Y_N^j - Y_N^{(0)}}{\sigma_\rho^2} \right) \quad \text{for } j = 1, 2, 3.$$ 

$G_j$ is expected to approximately follow the standard normal distribution under $H_0$ when the sample size is large.
The final test statistic of the study with the given interim selection rule is then:

\[ W = G_j, \text{ if } \bar{X}_{n_1}^{(j)} = \max \left( \bar{X}_{n_1}^{(1)}, \bar{X}_{n_1}^{(2)}, \bar{X}_{n_1}^{(3)} \right) \quad j = 1, 2, 3. \]

That is, conditional on the interim selection, \( W \) takes on the value of the effect from the “winner” treatment group as the final test statistic.

In the following, we derive the asymptotic distribution of the final test statistic. For the very general case under \( H_1 \), the distribution of the final test statistic \( W \) could be written as:

\[ F_W(w) = \sum_{j=1}^{3} P(P_{j1} < a_j, P_{j2} > b_j, P_{j3} > c_j) \]

where:

\[ P_{j1} = G_j - \sqrt{\frac{N}{\left[ (u_j^{X*})^2 \sigma_{Xj}^2 + (u_0^{X*})^2 \sigma_{T0}^2 + 2 \right]}} (u_j^{Y*} - u_0^{Y*}) \]

\[ P_{j2} = \frac{\bar{X}_{n_1}^{(j)}}{t} - \frac{\bar{X}_{n_1}^{(h)}}{t} - \left( \frac{u_j^{X*} \sigma_X}{t} - \frac{u_h^{X*} \sigma_X}{t} \right) \]

\[ P_{j3} = \frac{\bar{X}_{n_1}^{(j)}}{t} - \frac{\bar{X}_{n_1}^{(l)}}{t} - \left( \frac{u_j^{X*} \sigma_X}{t} - \frac{u_l^{X*} \sigma_X}{t} \right) \]

\[ a_j = w - \sqrt{\frac{N}{\left[ (u_j^{X*})^2 \sigma_{Xj}^2 + (u_0^{X*})^2 \sigma_{T0}^2 + 2 \right]}} (u_j^{Y*} - u_0^{Y*}) \]

\[ b_j = - \left( \frac{u_j^{X*} \sigma_X}{t} - \frac{u_h^{X*} \sigma_X}{t} \right) \]

\[ c_j = - \left( \frac{u_j^{X*} \sigma_X}{t} - \frac{u_l^{X*} \sigma_X}{t} \right) \]

\[ t = \frac{\sigma_X^2}{\sqrt{n_1}} \]

\[ j \neq h, j \neq l, h \neq l, j, h, l \in \{1, 2, 3\}. \]
\[(P_{j_1}, P_{j_2}, P_{j_3})'\] is approximately from a multi-normal distribution,

\[
\begin{pmatrix}
P_{j_1} \\
P_{j_2} \\
P_{j_3}
\end{pmatrix} \sim N\left(\begin{pmatrix}0 \\ Y_j \\ Y_j \end{pmatrix}, \begin{pmatrix}1 & Y_j & Y_j \\ Y_j & 2 & 1 \\ Y_j & 1 & 2 \end{pmatrix}\right)
\]

where \(Y_j = \sqrt{\frac{n_1 \left(\frac{\bar{u}_j^X}{\bar{\sigma}^2_{j_1} + 1}\right)}{N\left(\frac{\bar{u}^X}{\bar{\sigma}^2_{j_1} + \bar{\sigma}^2_{j_0} + 2}\right)} \rho} \).

Thus,

\[
F_w(w) = \sum_{j=1}^{3} \int_{-\infty}^{w} N\left(
\begin{pmatrix}
\bar{u}_j^X \\
\bar{u}_j^X
\end{pmatrix}, \begin{pmatrix}
\bar{\sigma}^2_{j_1} + 1 \\
\bar{\sigma}^2_{j_0} + 2
\end{pmatrix}\right)

\left(\begin{pmatrix}
\bar{u}_j^X - \bar{u}_j^X \\
\bar{u}_j^X - \bar{u}_j^X
\end{pmatrix} \begin{pmatrix}
1 & Y_j & Y_j \\ Y_j & 2 & 1 \\ Y_j & 1 & 2
\end{pmatrix} \begin{pmatrix}
\bar{u}_j^X - \bar{u}_j^X \\
\bar{u}_j^X - \bar{u}_j^X
\end{pmatrix}\right)

f_j(p_{j_1}, p_{j_2}, p_{j_3}) dp_{j_1} dp_{j_2} dp_{j_3}
\]

where:

\(j \neq h, j \neq l, h \neq j, h, l \in \{1, 2, 3\},\)

\(f_j(p_{j_1}, p_{j_2}, p_{j_3})\) is p.d.f. of \(N\left(\begin{pmatrix}0 \\ Y_j \\ Y_j \end{pmatrix}, \begin{pmatrix}1 & Y_j & Y_j \\ Y_j & 2 & 1 \\ Y_j & 1 & 2 \end{pmatrix}\right),\)

\[Y_j = \sqrt{\frac{n_1 \left(\frac{\bar{u}_j^X}{\bar{\sigma}^2_{j_1} + 1}\right)}{N\left(\frac{\bar{u}^X}{\bar{\sigma}^2_{j_1} + \bar{\sigma}^2_{j_0} + 2}\right)} \rho}.\]

### 3.4.4 Type I Error Control

It is shown that the interim treatment selection will skew the distribution of the final test statistic. Hence the type I error rate of the design might be inflated if no statistical adjustment is made. In this section, we find the adjusted critical rejection values of the final test statistic for preserving the type I error rate.
It is easy to show that, under $H_0$, the distribution of final test statistic is:

$$F_0(w) = \sum_{j=1}^{3} \int_{-\infty}^{w} \int_{0}^{+\infty} \int_{0}^{+\infty} f_j(p_{j1}, p_{j2}, p_{j3}) dp_{j3} dp_{j2} dp_{j1}$$

where:

$$f_j(p_{j1}, p_{j2}, p_{j3}) \text{ is p.d.f. of } N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \gamma_j & \gamma_j \\ \gamma_j & 2 & 1 \\ \gamma_j & 1 & 2 \end{pmatrix}\right).$$

$$\gamma_j = \sqrt{\frac{n_3 \left(\hat{u}_{0}\right)^2 \sigma_{\gamma_j}^2 + 1}{N\left[\left(\hat{u}_{0}\right)^2 \sigma_{\gamma_j}^2 + \left(\hat{u}_{0}\right)^2 \sigma_{\gamma_j}^2 + 2\right]}} \rho.$$ 

Let $w_\alpha$ be the upper $100\alpha$ percent quantile of $F_0$,

$$w_\alpha = F_W^{-1}(1 - \alpha|H_0) = F_0^{-1}(1 - \alpha)$$

The type I error rate of the design is controlled at level $\alpha$ if the 1-sided rejection region is

$$\Omega = \{W: W > w_\alpha\}.$$ 

The numerical values of $w_\alpha$ could be calculated using the R function \texttt{wang_cv_K4} developed by us in Chapter 2.

In addition, if we assume $\sigma_{\gamma_1}^2 = \sigma_{\gamma_2}^2 = \sigma_{\gamma_3}^2 = \sigma_{\gamma_0}^2$, that is the variability of the estimated mean-level correlation for each treatment group is the same, the distribution of final test statistic $W$ under $H_0$ can be approximated by:

$$F_0(w) = 3 \int_{-\infty}^{w} \int_{0}^{+\infty} \int_{0}^{+\infty} f(p_0, p_1, p_2) dp_0 dp_1 dp_2$$

where $f(p_0, p_1, p_2)$ is p.d.f. of $N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \gamma & \gamma \\ \gamma & 2 & 1 \\ \gamma & 1 & 2 \end{pmatrix}\right), \gamma = \sqrt{\frac{n_3}{2N}} \rho.$

The numerical values of $w_\alpha$ could be calculated using the R function \texttt{convention_cv_K4} in this case.
3.5 Performance Evaluation of Two-Stage Winner Design under Proposed Model

In this section, we show performance evaluation of the two-stage winner design using our proposed two-level correlation model. As illustrated earlier, when the rank order of mean responses of the biomarker is assumed to be the same as that of the primary endpoint, the power of a two-stage winner design could be easily overestimated if the conventional one-level correlation model is used. It will be shown that the new assessment approach is more reasonable, and can help with determining when a two-stage winner design should be used.

Assume \( u_0^X = 1, \sigma_x^2 = 1, u_0^Y = 1, \sigma_Y^2 = 1 \) – this could always be achieved by scaling – and assume \( u_1^Y = 1.1, u_2^Y = 1.5, u_3^Y = 1.3 \). We assume biomarker and primary endpoint are positively related, that is, large values of biomarker measurements correspond to large values of the primary endpoint. We consider the following cases when the means of biomarker and primary endpoint are correlated in linear and in non-linear ways (Figure 3.2).
3.5.1 \( u^{X*} \) and \( u^{Y*} \) are Linearly Related

If the standardized mean of biomarker \( u^{X*} \) and the standardized mean of primary endpoint \( u^{Y*} \) are linearly related, as shown in Figure 3.2, then the mean-level correlation \( r_j \) is the same for all treatment groups. Under our assumption here, \( r_1 = r_2 = r_3 = r_0 = 1 \). Therefore, \( u_1^{X*} = 1.1, u_2^{X*} = 1.5, u_3^{X*} = 1.3 \), in this case.

Consider the design with interim sample size \( n_1 = 52 \) and maximum sample size \( N = 104 \); and \( n_1 = 67 \) and \( N = 134 \); respectively. The two sample size will yield 80% and 90% power, respectively, of the corresponding classical design with no interim adaptation.

Table 3.2 and Table 3.3 list the simulation results of the two-stage winner design under our setting for the above sample sizes for different values of \( \rho \) and \( \sigma_r^2 \). All the simulation results are based on 10,000 simulation iterations.
We can see that, when $\sigma_r^2$ is fixed, the power of the design for different values of $\rho$ is similar. For example, in Table 3.2, when $\sigma_r^2 = 0.2$, the power of the design is around 79%~80%. However, when $\rho$ is fixed, the power of the design has a significant drop when $\sigma_r^2$ increases. For example, in Table 3.2, for $\rho = 0.5$, the power of the design drops from 88% to 79.4% when $\sigma_r^2$ increases from 0 to 0.2. The results indicate that the individual level correlation $\rho$ has only a little influence on the performance of the two-
stage winner design, while \( \sigma_F^2 \), which measures the uncertainty of the estimated mean-level correlation has a significant influence on the design performance. Therefore, it is necessary to consider and incorporate \( \sigma_F^2 \) when evaluating a biomarker informed adaptive design. In the conventional one-level correlation model, since \( \sigma_F^2 \) is not considered, the power of the design can be easily overestimated. The simulation results also suggest that the two-stage winner design is not necessarily always better than the corresponding classical design. In our setting here, only when \( \sigma_F^2 < 0.2 \) does the two-stage winner design show its advantage in terms of power.

3.5.2 \( u^X^* \) and \( u^Y^* \) are Not Linearly Related

Case 1 and case 2 in Figure 3.2 show two possible shapes when the mean of biomarker \( u^X^* \) and mean of primary endpoint \( u^Y^* \) are not linearly related.

We assume \( u_1^X^* = 1.2, \ u_2^X^* = 2.5, \ u_3^X^* = 1.8 \) in case 1 and \( u_1^X^* = 1.05, \ u_2^X^* = 1.25, \ u_3^X^* = 1.15 \) in case 2. Simulation results of the two-stage winner design with interim sample size \( n_1 = 52 \) and maximum sample size \( N = 104 \) for these two cases are listed in Table 3.4 and Table 3.5 respectively. The results agree with the previous findings when \( u^X^* \) and \( u^Y^* \) are linearly related. The individual level correlation \( \rho \) does not show large influence on the power performance of the design; however, \( \sigma_F^2 \) shows significant influence.
Table 3.4 Power when $u_X^*$ and $u_Y^*$ are Not Linearly Related – Case 1 ($n_1 = 52, N = 104$)

<table>
<thead>
<tr>
<th>Power</th>
<th>$\rho$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>$\sigma_r^2$</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0.1</td>
</tr>
</tbody>
</table>

Table 3.5 Power when $u_X^*$ and $u_Y^*$ are Not Linearly Related – Case 2 ($n_1 = 52, N = 104$)

<table>
<thead>
<tr>
<th>Power</th>
<th>$\rho$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>$\sigma_r^2$</td>
<td>0</td>
</tr>
</tbody>
</table>

Different from what has been shown when $u_X^*$ and $u_Y^*$ are linearly related, under the setting for case 1, the simulation results suggest that the two-stage winner design is a better option than the corresponding classical design only when $\sigma_r^2 < 0.1$; the simulation results also suggest that, under the setting for case 2, the two-stage winner design would never be better than the corresponding classical design. Therefore, the two-stage winner design is not necessarily better than the corresponding classical design. The performance of the two-stage winner design is affected by how $u_X^*$ and $u_Y^*$ are related and the uncertainty about the mean-level correlation.

Another interesting index for the performance of the two-stage winner design is “power with best treatment”, which is the probability that the final hypothesis will be rejected
when the best treatment is selected at interim. Table 3.6 – Table 3.9 list the “power with best treatment” under the above scenarios associated with Tables 3.2 – Table 3.5.

Table 3.6 Power with Best Treatment when $u^{X_1}$ and $u^{X_2}$ are Linearly Related ($n_1 = 52, N = 104$)

<table>
<thead>
<tr>
<th>Power</th>
<th>$\rho$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>$\sigma^2_r$</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>0.2</td>
</tr>
</tbody>
</table>

Table 3.7 Power with Best Treatment when $u^{X_1}$ and $u^{X_2}$ are Linearly Related ($n_1 = 67, N = 134$)

<table>
<thead>
<tr>
<th>Power</th>
<th>$\rho$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>$\sigma^2_r$</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>0.2</td>
</tr>
</tbody>
</table>
Table 3.8 Power with Best Treatment when $u^x$ and $u^y$ are Not Linearly Related – Case 1 ($n_1 = 52, N = 104$)

<table>
<thead>
<tr>
<th>$\sigma_r^2$</th>
<th>Power</th>
<th>$\rho$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>0.2</td>
</tr>
<tr>
<td>0</td>
<td>90%</td>
<td>90%</td>
</tr>
<tr>
<td>0.1</td>
<td>80%</td>
<td>81%</td>
</tr>
</tbody>
</table>

Table 3.9 Power with Best Treatment when $u^x$ and $u^y$ are Not Linearly Related – Case 2 ($n_1 = 52, N = 104$)

<table>
<thead>
<tr>
<th>$\sigma_r^2$</th>
<th>Power</th>
<th>$\rho$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>0.2</td>
</tr>
<tr>
<td>0</td>
<td>60%</td>
<td>60%</td>
</tr>
</tbody>
</table>

It has been seen that, in general, the “power with best treatment” is lower than the power. However, the difference is affected by the shape of the mean-level relationship. When $u^x$ and $u^y$ are not linearly related, and have a relationship shape similar with the case 1 in Figure 3.2, the difference between “power with best treatment” and the power is small, as can be seen by comparing Table 3.4 and Table 3.8. When $u^x$ and $u^y$ are not linearly related, and have a relationship shape similar to case 2 in Figure 3.2, the difference between the two powers is large. These facts indicate that the probability of choosing the best treatment at interim is affected by the shape of the mean-level relationship of the two endpoints used in design.
3.6 Parameter Estimation

In this section, we propose a solution to the question “How to estimate the parameters incorporated in the two-level correlation model using historical data?”

It is clear that the two-level correlation model incorporates the following parameters: $\sigma_X$, $u_j^{X*}$, $\sigma_Y$, $u_j^{Y*}$, $\sigma_{rj}^2$ and $\rho$, $j = 0, \ldots, K - 1$.

Assume there are $n_j$ pairs of historical data for treatment $j$ on biomarker and primary endpoint, $j = 0, \ldots, K - 1$. Let $s_{x,j}^2$ be sample variance of biomarker $X_{i(j)}$ in treatment group $j$, $i = 1, \ldots, n_j$, and $\bar{x}^{(j)}$ be the observed sample mean of biomarker in treatment group $j$. Let $s_{y,j}^2$ be sample variance of primary endpoint $Y_{i(j)}$ in treatment group $j$, and $\bar{y}^{(j)}$ be the observed sample mean of primary endpoint in treatment group $j$.

We suggest estimating the parameters in the following natural way:

$$\hat{\sigma}_X^2 = \frac{\sum(n_j s_{x,j}^2)}{\sum n_j} \cdot \hat{\sigma}_X = \sqrt{\hat{\sigma}_X^2}$$

$$u_j^{X*} = \frac{\bar{x}^{(j)}}{\hat{\sigma}_X}$$

$$\hat{\sigma}_Y^2 = \frac{\sum(n_j s_{y,j}^2)}{\sum n_j} \cdot \hat{\sigma}_Y = \sqrt{\hat{\sigma}_Y^2}$$

$$u_j^{Y*} = \frac{\bar{y}^{(j)}}{\hat{\sigma}_Y}$$

$$\hat{\sigma}_{rj}^2 = \frac{(u_j^{Y*})^2}{(u_j^{X*})^2 + \frac{1}{n_j}} \left[ \frac{1}{n_j} \hat{\sigma}_Y^2 + \frac{1}{n_j} \hat{\sigma}_X^2 - 2 \frac{1}{n_j} \hat{\rho} - \frac{2}{n_j} \frac{1}{n_j} \right]$$

$$\hat{\rho} = Corr(X, Y)$$

which is the observed correlation coefficient of the pooled sample of $X_{i(j)}$ and $Y_{i(j)}$, $i = 1, \ldots, n_j, j = 0, 1, \ldots, K - 1$. 

(3.1)
Detailed derivation of Equation (3.1) is provided in Appendix B. An approximation for the variance of ratio was used in derivation (Stuart and Ord, 1998).

There are times when a common variance is preferred for $R_j$, i.e., $\sigma_{r_1}^2 = \cdots = \sigma_{r_{K-1}}^2 = \sigma_{r_0}^2 = \sigma_r^2$ will be assumed. For this case, we suggest: $\hat{\sigma}_r^2 = \frac{\sum (n_j \hat{\sigma}_{r_j}^2)}{\sum n_j}$.

Table 3.10 lists the simulation results, and compares the true parameter values with their estimates. The 3 cases in Figure 3.2 were considered. We assume $\rho = 0.5$, and the simulation results are based on 50 pairs of historical data for each treatment group. The numbers show that the proposed approach provides reasonable estimation for parameters.
Table 3.10 Estimation of the Parameters

<table>
<thead>
<tr>
<th></th>
<th>Linear</th>
<th>Nonlinear 1</th>
<th>Nonlinear 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>True</td>
<td>Estimate</td>
<td>True</td>
</tr>
<tr>
<td>( \rho )</td>
<td>0.5</td>
<td>0.51</td>
<td>0.5</td>
</tr>
<tr>
<td>( \sigma^2_y )</td>
<td>1</td>
<td>1.00</td>
<td>1</td>
</tr>
<tr>
<td>( u^y_0 )</td>
<td>1</td>
<td>1.01</td>
<td>1</td>
</tr>
<tr>
<td>( u^y_1 )</td>
<td>1.1</td>
<td>1.11</td>
<td>1.1</td>
</tr>
<tr>
<td>( u^y_2 )</td>
<td>1.5</td>
<td>1.50</td>
<td>1.5</td>
</tr>
<tr>
<td>( u^y_3 )</td>
<td>1.3</td>
<td>1.31</td>
<td>1.3</td>
</tr>
<tr>
<td>( \sigma^2_x )</td>
<td>1</td>
<td>1.00</td>
<td>1</td>
</tr>
<tr>
<td>( u^x_0 )</td>
<td>1</td>
<td>1.00</td>
<td>1</td>
</tr>
<tr>
<td>( u^x_1 )</td>
<td>1.1</td>
<td>1.11</td>
<td>1.2</td>
</tr>
<tr>
<td>( u^x_2 )</td>
<td>1.5</td>
<td>1.51</td>
<td>2.5</td>
</tr>
<tr>
<td>( u^x_3 )</td>
<td>1.3</td>
<td>1.30</td>
<td>1.8</td>
</tr>
<tr>
<td>( r^0 )</td>
<td>1</td>
<td>1.02</td>
<td>1</td>
</tr>
<tr>
<td>( r^1 )</td>
<td>1</td>
<td>1.01</td>
<td>0.92</td>
</tr>
<tr>
<td>( r^2 )</td>
<td>1</td>
<td>1.00</td>
<td>0.6</td>
</tr>
<tr>
<td>( r^3 )</td>
<td>1</td>
<td>1.01</td>
<td>0.72</td>
</tr>
<tr>
<td>( \sigma^2_r )</td>
<td>0.014</td>
<td>0.015</td>
<td>0.01</td>
</tr>
</tbody>
</table>
3.7 Discussion

In this new age of personalized medicine, biomarker informed adaptive designs are very attractive. They have the potential to shorten the time of drug development and to bring the right drug to patient earlier, by incorporating correlated short-term endpoint at the interim stages.

It is strongly suggested that statistical simulations should be performed before a drug trial to understand the operating characteristics of the trial design. As an important component of this process, a model need to be specified for simulation of the two endpoints incorporated in the study. The conventional one-level correlation model used might be inappropriate when the relationship between the biomarker and the primary endpoint is not well known. This model only considers the individual-level correlation between the interim and final endpoint of the design, but ignores the uncertainty of the estimated mean-level correlation between the two endpoints. Hence, the simulation results of a biomarker informed adaptive design using the conventional one-level model can easily misestimate the power.

We propose a new two-level correlation model for fitting the two endpoints in a biomarker informed adaptive design in this chapter. The new two-level correlation model incorporates correlations at both the individual and mean level. It is shown that in a biomarker informed adaptive design, the power is much more sensitive to the correlation between biomarker and primary endpoint at mean level than the correlation at the individual level. Simulations using the two-level correlation model for biomarker informed designs produce more sensible and reasonable results.
The proposed two-level correlation model is illustrated in the context of a two-stage winner design in this chapter. With the derived distribution of the test statistics and stopping boundary information, the type I error rate can be controlled. We considered 3 cases where the mean-level correlations are in different shapes. It is shown that the shape, together with the uncertainty about the shape, should be counted when comparing the biomarker informed adaptive design with the corresponding classical design. An absolute advantage of the biomarker informed design is not guaranteed. In addition, it's also shown that the shape of the mean level correlation affects the probability of choosing the best treatment at interim of the design.

Methods were proposed to estimate the parameters. The proposed estimators for the parameters appear to be unbiased by simulations. Hence, based on the prior historical data, we will be able to answer questions such as “which design will provide higher power, the biomarker informed adaptive design or the classical Dunnett design”, and, “what should the sample size be in order to get 80% power” by simulations.

In general, when a good portion of the relationship between the biomarker and the primary endpoint is known, the biomarker informed design is recommended.
CHAPTER 4
BIOMARKER INFORMED ADD-ARM DESIGN FOR UNIMODAL RESPONSE

4.1 Introduction

In drug development, identification of effective doses (or therapies) and inference on selected treatments are usually performed in two separate trials, i.e. separate phase II and phase III trials.

For dose-finding, the traditional approach relies on analysis of variance (ANOVA) to compare the efficacy of each dose to control, using Dunnett’s adjustment (Dunnett, 1955) for multiplicity adjustment. The data analysis is pre-specified and no adaptation is used in either the design or the analysis. Adaptive approaches have also been proposed for designing and analyzing dose-response studies, with the purpose of optimizing the assessment of the dose-response relationship and narrowing down the set of candidate doses to select the estimated target dose. In general, these adaptive procedures can be categorized into two classes: “analysis-focused adaptive approaches” and “design-focused adaptive approaches” (Bornkamp et al. 2007).

The “analysis-focused adaptive approaches” relies on a fixed study design, that is, no design modifications are performed during the study. In this case, the focus is on selecting the most appropriate data analysis method. The choice of the “best” method is adaptive in the sense that it is driven by the data collected in the trial. The multiple comparison procedures-modeling approach (MCP-Mod) proposed by Bretz et al. (2005), Bayesian modeling-averaging approach (BMA), multiple trend test approach (MTT) and
nonparametric dose-response modeling approach (LOCFIT) are analysis-focused adaptive approaches.

The “design-focused adaptive approaches” include methods that allow modifications of certain elements of the study design based on the data collected in the trial. For example, based on the review of interim safety or efficacy data, the trial’s sponsor can decide to drop one or more doses or reduce the number of patients assigned to these doses. This adaptation can be performed in a continuous manner, i.e., the design can be updated prior to the enrollment of each new patient. Alternatively, one can consider group-sequential adaptive strategies and perform design modifications based on responses from cohorts of patients. Examples of design-focused adaptive approaches include: Bayesian general adaptive dose allocation (GADA) and D-optimal response-adaptive approach (D-Opt).

Recently, adaptive seamless designs have been explored. These designs combine dose selection and confirmation into one trial, so that the overall development time for a drug might be shortened and the number of patients needed for the trial might be reduced. A well-known example of the seamless design is Phase II/III drop-the-loser design (Bauer et al. 1999). Chang and Wang (2013) proposed a 3-stage add-arm design for unimodal dose-response, which is also an example of seamless design for dose selection and confirmation. This design utilizes the prior knowledge about the dose-response to optimize the procedure of dose-finding. If it is known that the dose-response curve is unimodal (e.g. anti-angiogenic agents or mixed agonist-antagonists), some poor candidate doses could be identified based on interim outcomes and hence patients need not be
exposed to those doses. It was shown that this design can save sample size by as much as 20% compared with two-stage drop-the-loser design with unimodal response.

One potential challenge of both the drop-the-loser design and the 3-stage add-arm design for unimodal response is that, the benefits of the designs might be limited if the endpoint at interim analysis takes very long to observe. For example in the 3-stage add-arm design: the arm(s) added at the second stage are determined by the observed data from the first stage. An endpoint that takes long to observe at interim can be operationally challenging in this case. Besides, the long “white space” between the two stages might add a temporal effect.

To circumvent some of these issues, we propose a “biomarker informed add-arm design for unimodal response” in this manuscript. The new design contributes to optimizing the procedure of dose-finding when a biomarker of the study primary endpoint exists and prior evidence indicates a unimodal dose-response relationship. In our proposed design, the interim decisions are based on the measurements of the biomarker, which is a short-term endpoint that is correlated with the primary endpoint. Without loss of generality and for the purpose of this discussion, we assume the biomarker and the primary endpoint are positively related, that is, large values of biomarker measurement correspond to large values of the primary endpoint. Designs with up to 7 active treatment arms are considered. For the two endpoints (the biomarker and primary endpoint) incorporated in the design, we use the two-level correlation model developed in Wang et al. (2013) to fit them. We propose a statistical approach that controls the type I error rate of the proposed design, and carry out extensive simulation studies to study the power performance of the
design. The new design is compared against its corresponding biomarker informed two-stage winner design (Shun et al. 2008) in a variety of settings, including different relationship between the two endpoints and different unimodal shapes of dose-response curve. Our proposed design is shown to outperform the corresponding biomarker informed two-stage winner design in power on an average.

4.2 Biomarker Informed Add-Arm Design with 4 Active Treatment Arms

4.2.1 The 3-Stage Design

We make the assumption that the dose-response curve for the candidate drug is unimodal, and there exists a biomarker which is positively (or negatively) correlated with the response. In this section, we present the 3-stage biomarker informed add-arm design we proposed.
Assume there are 4 active treatment arms (dose levels 1, 2, 3, 4) and a control arm. The order of dose levels suggests an increasing potency. The biomarker informed add-arm design starts with randomizing patients into mid-level doses 2, 3, and the control group (denoted by “dose 0”). After $n_1$ patients are treated with either dose 2 or dose 3, the observed measurements of biomarker for the two treatment groups are compared. If dose 2 group has better response on biomarker than dose 3 group, patients are randomized to dose 1 group and the control group for the next stage. Likewise, if dose 3 group has better response on biomarker than dose 2 group, patients are randomized to dose 4 group and the control group at stage two.

At the second interim look, the dose group added after evaluation at stage one will be compared with the “winner” dose group from stage one on biomarker. For example, if dose 2 is the “winner” from stage one, then dose 1 will be added into the study, hence at
the second interim analysis the biomarker results for dose 1 group will be compared to that for dose 2 group. The winner at this stage will be selected as the “best” dose candidate to carry to the end of the study with the control group. Additional \( n_2 \) patients will be used for both of the “best” dose candidate and the control, and the final analysis will use patients from both stages on study primary endpoint. In this design, patients need not be exposed to every candidate dose.

4.2.2 Mathematical Notations for the Design

Let \( K \) be number of treatment groups (\( K - 1 \) active treatment groups, and 1 control group), \( N \) be the maximum sample size for each dose, \( N = n_1 + n_2 \). Let \( \{X^{(j)}_i|i = 1, ..., n_1\} \) be measurements of the biomarker for \( i \)th patient in \( j \)th dose group; and \( \{Y^{(j)}_i|i = 1, ..., N\} \) be measurements of the study primary endpoint. \( j = 0, 1, ..., K - 1 \). \( j = 0 \) represents the control group and \( j = 1, ..., K - 1 \) the active treatment groups. For simplicity, assume both the endpoints are normally distributed.

We use the two-level correlation model proposed in Wang et al. (2013) for modeling the two endpoints, that is, assume

\[
\begin{pmatrix}
\frac{X^{(j)}_i}{\sigma_X} \\
\frac{Y^{(j)}_i}{\sigma_Y} | R_j
\end{pmatrix} 
\sim N\left(\begin{pmatrix} u_j^{x^*} \\ R_j u_j^{x^*} \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right)
\]

\( R_j \sim N\left(r_j, \sigma_r^2\right) \)
where $\sigma^2_X$ is the variance of biomarker, $\sigma^2_Y$ is the variance of study primary endpoint, $u^X_j$ is the standardized mean of biomarker for treatment group $j$. $R_j$ is a random variable that describes the observed mean level correlation (or “mean level association”) between the biomarker and primary endpoint for treatment group $j$ from historical data. Assume $R_j$ is normally distributed, with mean $r_j$, which is the true mean level correlation between the biomarker and the primary endpoint; and variance $\sigma^2_R$. The standardized mean of study primary endpoint for treatment group $j$, $u^Y_j$, can therefore be expressed as $u^Y_j = r_j u^X_j$.

Assume $\{X^{(j)}_i | i = 1, ..., n_1\}$ and $\{Y^{(j)}_i | i = 1, ..., n_1\}$ are correlated with a correlation $\rho$ for the same $j$ and $i$.

Let $\bar{X}^{(j)}_{n_1} = \frac{1}{n_1} \sum X^{(j)}_i$ be the mean of biomarker measurements for treatment group $j$ at interim, and $\bar{Y}^{(j)}_N = \frac{1}{N} \sum Y^{(j)}_i$ be the mean of primary endpoint measurements for treatment group $j$ at final. The detail randomization and dose selection procedures of the design are as follows (Figure 4.2):
Stage 1: Assign $2.5n_1$ subjects in dose group 2, 0, and 3 using randomization ratio 1:0.5:1.

Stage 2: If the observations $\overline{x}_{n_1}^{(2)} > \overline{x}_{n_1}^{(3)}$, assign $1.5n_1$ subjects in dose group 0 and 1 using 0.5:1 randomization. Otherwise, if the observations $\overline{x}_{n_1}^{(2)} \leq \overline{x}_{n_1}^{(3)}$, assign $1.5n_1$ subjects in dose group 0 and 4 using 0.5:1 randomization.

Stage 3: (a) If $\overline{x}_{n_1}^{(2)} > \overline{x}_{n_1}^{(3)}$ and $\overline{x}_{n_1}^{(2)} \geq \overline{x}_{n_1}^{(1)} + c_R$, select dose 2 as the best dose candidate among the 4 dose candidates; otherwise, select dose 1 as the best dose candidate. If $\overline{x}_{n_1}^{(2)} \leq \overline{x}_{n_1}^{(3)}$ and $\overline{x}_{n_1}^{(3)} \geq \overline{x}_{n_1}^{(4)} + c_R$, select dose 3 as the best dose candidate among the 4 dose candidates; otherwise, select dose 4 as the best dose candidate. (b) Assign $2n_2$
subjects in the selected best candidate dose group and the control group using 1:1 randomization.

In total, there will be $4n_1 + 2n_2$ subjects. The final analysis will use measurements of study primary endpoint from $n_1 + n_2$ subjects in the selected best dose candidate group and the control group.

Notice that the randomization ratio being used at the first stage of the design is 1:0.5:1 instead of 1:1:1. This is because that 1:0.5:1 can keep the treatment blinding and balance the confounding factors at the second stage. If 1:1:1 was used, then $n_1$ subjects would be randomized to the control group at the first stage, hence all the subjects have to be assigned to the active treatment group without randomizing at the second stage, thus unblinding the treatment and potentially imbalance of any baseline confounding factors.

Also notice that the criteria being used at the second stage for selecting the best candidate dose is comparing $\bar{x}_{n_1}^{(2)}$ with $\bar{x}_{n_1}^{(1)} + c_R$, and comparing $\bar{x}_{n_1}^{(3)}$ with $\bar{x}_{n_1}^{(4)} + c_R$. In next section we discuss how to determine the constant $c_R$.

4.2.3 Find $c_R$

Define the term “selection probability” as the probability of selecting a dose level as the preferred treatment for the next stage of the adaptive trial.

If $c_R = 0$, the proposed design becomes a biomarker informed 3-stage winner design, in which the active treatment group with the maximum observed response on biomarker at the current stage is picked as the winner. The problem in this approach is the selection
probability will be skewed for a flat response curve. More specifically in our case, the
selection probability of dose 1, 2, 3, and 4 will be 1/6, 2/6, 2/6 and 1/6, respectively when
in fact all doses have the same effect.

We expect equal selection probability for each dose group while all doses have the same
effect. Thus, we force the probability of selecting each dose equal (or at least
approximately equal) under $H'_0$: $u_1^* = u_2^* = u_3^* = u_4^*$, i.e., the following equations
need to hold (or approximately hold):

$$P\left( \frac{X_{n_1}^{(2)}}{\sqrt{\frac{2\sigma_X}{n_1}}} > \frac{X_{n_1}^{(3)}}{\sqrt{\frac{2\sigma_X}{n_1}}} , \frac{X_{n_1}^{(2)}}{\sqrt{\frac{2\sigma_X}{n_1}}} \geq \frac{X_{n_1}^{(1)}}{\sqrt{\frac{2\sigma_X}{n_1}}} + c_R ; H'_0 \right) = \frac{1}{4}$$

$$P\left( \frac{X_{n_1}^{(2)}}{\sqrt{\frac{2\sigma_X}{n_1}}} > \frac{X_{n_1}^{(3)}}{\sqrt{\frac{2\sigma_X}{n_1}}} , \frac{X_{n_1}^{(2)}}{\sqrt{\frac{2\sigma_X}{n_1}}} < \frac{X_{n_1}^{(1)}}{\sqrt{\frac{2\sigma_X}{n_1}}} + c_R ; H'_0 \right) = \frac{1}{4}$$

$$P\left( \frac{X_{n_1}^{(2)}}{\sqrt{\frac{2\sigma_X}{n_1}}} \leq \frac{X_{n_1}^{(3)}}{\sqrt{\frac{2\sigma_X}{n_1}}} , \frac{X_{n_1}^{(3)}}{\sqrt{\frac{2\sigma_X}{n_1}}} \geq \frac{X_{n_1}^{(4)}}{\sqrt{\frac{2\sigma_X}{n_1}}} + c_R ; H'_0 \right) = \frac{1}{4}$$

$$P\left( \frac{X_{n_1}^{(2)}}{\sqrt{\frac{2\sigma_X}{n_1}}} \leq \frac{X_{n_1}^{(3)}}{\sqrt{\frac{2\sigma_X}{n_1}}} , \frac{X_{n_1}^{(3)}}{\sqrt{\frac{2\sigma_X}{n_1}}} < \frac{X_{n_1}^{(4)}}{\sqrt{\frac{2\sigma_X}{n_1}}} + c_R ; H'_0 \right) = \frac{1}{4}$$

Let $D_1 = \frac{X_{n_1}^{(2)} - X_{n_1}^{(1)}}{\sqrt{\frac{2\sigma_X}{n_1}}}$, $D_2 = \frac{X_{n_1}^{(2)} - X_{n_1}^{(1)}}{\sqrt{\frac{2\sigma_X}{n_1}}}$, the conditional probability can be written as:

$$P\left( \frac{X_{n_1}^{(2)}}{\sqrt{\frac{2\sigma_X}{n_1}}} > \frac{X_{n_1}^{(3)}}{\sqrt{\frac{2\sigma_X}{n_1}}} , \frac{X_{n_1}^{(2)}}{\sqrt{\frac{2\sigma_X}{n_1}}} \geq \frac{X_{n_1}^{(1)}}{\sqrt{\frac{2\sigma_X}{n_1}}} + c_R ; H'_0 \right) = P\left( D_1 > 0 , D_2 \geq \frac{c_R}{\sqrt{\frac{2\sigma_X}{n_1}}} \right)$$

$$= \int_{0}^{+\infty} \int_{0}^{+\infty} f(d_1, d_2) d_2 d_1$$

where $(D_1, D_2) \sim N\left( \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix} \right)$ under $H'_0$, $f(d_1, d_2)$ is p.d.f. of $N\left( \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix} \right)$. 
The integration could be solved by the numerical software. It shows that \( \frac{c_R}{\sqrt{\frac{2\sigma_X^2}{n_1}}} \) is a constant, \( c_R = 0.39 \).

Thus, \( c_R \) is a function of \( n_1 \) and \( \sigma_X^2 \), \( c_R = 0.39 \sqrt{\frac{2\sigma_X^2}{n_1}} \).

\( \sigma_X^2 \) is suggested to be substituted by the sample variance \( \overline{\sigma_X^2} \) for calculating \( c_R \), as \( \sigma_X^2 \) is unknown in practice. The obtained \( c_R \) is suggested to be further validated via simulations to ensure an (approximately) equal selection probability of each dose group for a flat response curve.

4.2.4 Type I Error Control

The type I error rate of a trial that uses biomarker informed add-arm design can be preserved by adjusting the critical rejection value of its final test statistic. In this section, we demonstrate how to adjust the critical rejection value for type I error control and develop R function for this purpose.

Let us consider the following hypotheses:

\[ H_0: u_1^{Y*} = u_2^{Y*} = u_3^{Y*} = u_4^{Y*} = u_0^{Y*} \]

\[ H_1: u_1^{Y*} > u_0^{Y*} \text{ or } u_2^{Y*} > u_0^{Y*} \text{ or } u_3^{Y*} > u_0^{Y*} \text{ or } u_4^{Y*} > u_0^{Y*} \]

It is reasonable to assume that \( u_1^{Y*} = u_2^{Y*} = u_3^{Y*} = u_4^{Y*} = u_0^{Y*} \) when \( u_1^{Y*} = u_2^{Y*} = u_3^{Y*} = u_4^{Y*} = u_0^{Y*} \) and \( \rho \neq 0 \).
Let $G_j$ be the test statistic comparing the primary endpoint of the $j^{th}$ treatment group and the control group. The final test statistic for the biomarker informed add-arm design with 4 active treatments can then be expressed as the following:

$$W = \begin{cases} 
G_1 & \text{if } \overline{X}_{n_1}^{(2)} > \overline{X}_{n_1}^{(3)} \text{ and } \overline{X}_{n_1}^{(2)} < \overline{X}_{n_1}^{(1)} + c_R \\
G_2 & \text{if } \overline{X}_{n_1}^{(2)} > \overline{X}_{n_1}^{(3)} \text{ and } \overline{X}_{n_1}^{(2)} \geq \overline{X}_{n_1}^{(1)} + c_R \\
G_3 & \text{if } \overline{X}_{n_1}^{(2)} \leq \overline{X}_{n_1}^{(3)} \text{ and } \overline{X}_{n_1}^{(3)} \geq \overline{X}_{n_1}^{(4)} + c_R \\
G_4 & \text{if } \overline{X}_{n_1}^{(2)} \leq \overline{X}_{n_1}^{(3)} \text{ and } \overline{X}_{n_1}^{(3)} < \overline{X}_{n_1}^{(4)} + c_R
\end{cases}$$

where $G_j = \sqrt{\frac{N}{\sigma_j^2(u_j^{X_j})^2 + (u_0^{X_0})^2 + 2}} \left(\overline{Y}_N^{(j)} - \overline{Y}_N^{(0)}\right). j = 1, 2, 3, 4.$

That is, conditional on the interim selection, $W$ takes the value of the effect of the selected best dose candidate group as the final test statistic.

For a general case under $H_1$, the distribution of the final test statistic $W$ can be derived as the following:

$$F_W(w) = P(W < w)$$

$$= P\left(G_1 < w, \overline{X}_{n_1}^{(2)} > \overline{X}_{n_1}^{(3)}, \overline{X}_{n_1}^{(2)} < \overline{X}_{n_1}^{(1)} + c_R\right) + P\left(G_2 < w, \overline{X}_{n_1}^{(2)} > \overline{X}_{n_1}^{(3)}, \overline{X}_{n_1}^{(2)} \geq \overline{X}_{n_1}^{(1)} + c_R\right) + P\left(G_3 < w, \overline{X}_{n_1}^{(2)} \leq \overline{X}_{n_1}^{(3)}, \overline{X}_{n_1}^{(3)} \geq \overline{X}_{n_1}^{(4)} + c_R\right) + P\left(G_4 < w, \overline{X}_{n_1}^{(2)} \leq \overline{X}_{n_1}^{(3)}, \overline{X}_{n_1}^{(3)} < \overline{X}_{n_1}^{(4)} + c_R\right)$$

$$= P_{P_{11} < w} - \sqrt{\frac{N}{\left(\overline{u}_1^{x_1} \right)^2 \sigma_1^2 + \left(\overline{u}_0^{x_0}\right)^2 \sigma_1^2 + 2}} \left(u_1^{Y_1} - u_0^{Y_0}\right) P_{2} > \overline{u}_2^{x_2} - \overline{u}_3^{x_3} P_3 > \frac{c_R}{\sqrt{\overline{n}}} - \frac{\overline{u}_2^{x_2} - \overline{u}_1^{x_1}}{\sqrt{\frac{2}{\overline{n}}}}$$

$$+ P_{P_{21} < w} - \sqrt{\frac{N}{\left(\overline{u}_2^{x_2} \right)^2 \sigma_2^2 + \left(\overline{u}_0^{x_0}\right)^2 \sigma_2^2 + 2}} \left(u_1^{Y_1} - u_0^{Y_0}\right) P_{2} > \overline{u}_2^{x_2} - \overline{u}_3^{x_3} P_3 > \frac{c_R}{\sqrt{\overline{n}}} - \frac{\overline{u}_2^{x_2} - \overline{u}_1^{x_1}}{\sqrt{\frac{2}{\overline{n}}}}$$
\[ P( P_{31} < w - \sqrt{\frac{N}{\left( \frac{u_3^{Y*} - u_0^{Y*}}{\sigma_3^2} \right)^2 + \frac{\sigma_3^2}{2}} \cdot \frac{\sigma_3^2}{2} \cdot \frac{u_3^{Y*} - u_0^{Y*}}{\sigma_3^2} \cdot \frac{u_3^{Y*} - u_0^{Y*}}{\sigma_3} \cdot \frac{\sigma_3}{2}}})\]

\[ + P( P_{41} < w - \sqrt{\frac{N}{\left( \frac{u_4^{Y*} - u_0^{Y*}}{\sigma_4^2} \right)^2 + \frac{\sigma_4^2}{2}} \cdot \frac{\sigma_4^2}{2} \cdot \frac{u_4^{Y*} - u_0^{Y*}}{\sigma_4^2} \cdot \frac{u_4^{Y*} - u_0^{Y*}}{\sigma_4} \cdot \frac{\sigma_4}{2}}})\]

where:

\[ P_{j1} = \frac{G_j - \sqrt{\frac{N}{\left( \frac{u_j^{Y*} - u_0^{Y*}}{\sigma_j^2} \right)^2 + \frac{\sigma_j^2}{2}} \cdot \frac{\sigma_j^2}{2} \cdot \frac{u_j^{Y*} - u_0^{Y*}}{\sigma_j^2} \cdot \frac{u_j^{Y*} - u_0^{Y*}}{\sigma_j} \cdot \frac{\sigma_j}{2}}}}{\sqrt{\frac{2}{\sigma_j^2}} \cdot \frac{2}{\sigma_j} \cdot \frac{2}{\sigma_j} \cdot \frac{2}{\sigma_j} \cdot \frac{2}{\sigma_j}}\]

\[ P_2 = \frac{\frac{X_{n1}^{(2)} - X_{n1}^{(3)}}{2\sigma_2^2}}{\frac{2}{\sigma_2^2}} \cdot \frac{u_2^{Y*} - u_0^{Y*}}{\sqrt{\frac{2}{\sigma_2^2}} \cdot \frac{2}{\sigma_2} \cdot \frac{2}{\sigma_2} \cdot \frac{2}{\sigma_2} \cdot \frac{2}{\sigma_2}}\]

\[ P_3 = \frac{\frac{X_{n1}^{(2)} - X_{n1}^{(1)}}{2\sigma_3^2}}{\frac{2}{\sigma_3^2}} \cdot \frac{u_3^{Y*} - u_0^{Y*}}{\sqrt{\frac{2}{\sigma_3^2}} \cdot \frac{2}{\sigma_3} \cdot \frac{2}{\sigma_3} \cdot \frac{2}{\sigma_3} \cdot \frac{2}{\sigma_3}}\]

\[ P_4 = \frac{\frac{X_{n1}^{(3)} - X_{n1}^{(4)}}{2\sigma_4^2}}{\frac{2}{\sigma_4^2}} \cdot \frac{u_4^{Y*} - u_0^{Y*}}{\sqrt{\frac{2}{\sigma_4^2}} \cdot \frac{2}{\sigma_4} \cdot \frac{2}{\sigma_4} \cdot \frac{2}{\sigma_4} \cdot \frac{2}{\sigma_4}}\]

\( (P_{11}, P_{21}, P_3, (P_{21}, P_2, P_3), (P_{31}, P_2, P_4), (P_{41}, P_2, P_4) \)' are approximately multivariate normal.

\[
\begin{pmatrix}
P_{11} \\
P_2 \\
P_3
\end{pmatrix} \sim N
\begin{pmatrix}
0 \\
0 \\
0
\end{pmatrix},
\begin{pmatrix}
1 & 0 & -\lambda_1 \\
0 & 1 & \frac{1}{2} \\
-\lambda_1 & \frac{1}{2} & 1
\end{pmatrix}
\]
\[
\begin{pmatrix}
P_{21} \\
P_2 \\
P_3
\end{pmatrix}
\sim N
\begin{pmatrix}
0, \\
\begin{pmatrix}
1 & \lambda_2 \\
\lambda_2 & 1
\end{pmatrix} \\
\begin{pmatrix}
\lambda_2 \\
\frac{1}{2}
\end{pmatrix}
\end{pmatrix}
\]

\[
\begin{pmatrix}
P_{31} \\
P_2 \\
P_4
\end{pmatrix}
\sim N
\begin{pmatrix}
0, \\
\begin{pmatrix}
1 & -\lambda_3 \\
-\lambda_3 & 1
\end{pmatrix} \\
\begin{pmatrix}
-\lambda_3 \\
\frac{1}{2}
\end{pmatrix}
\end{pmatrix}
\]

\[
\begin{pmatrix}
P_{41} \\
P_2 \\
P_4
\end{pmatrix}
\sim N
\begin{pmatrix}
0, \\
\begin{pmatrix}
1 & 0 & -\lambda_4 \\
0 & 1 & -\frac{1}{2}
\end{pmatrix} \\
\begin{pmatrix}
-\lambda_4 \\
\frac{1}{2}
\end{pmatrix}
\end{pmatrix}
\]

where \( \lambda_j = \rho \sqrt{\frac{n_1}{2N}} \sqrt{\frac{(\bar{x}_j \cdot \sigma_j^2 + 1)}{(u_j \cdot \sigma_j^2 + u_0 \cdot \sigma_0^2 + 2)}}, \ j = 1, 2, 3, 4. \)

Under \( H_0 \), the distribution of the final test statistic can be simplified as:

\[
F_0(w) = P(W < w; H_0)
\]

\[
= P(P_{11} < w, P_2 > 0, P_3 < 0.39) + P(P_{21} < w, P_2 > 0, P_3 > 0.39)
\]

\[
+ P(P_{31} < w, P_2 \leq 0, P_4 \geq 0.39) + P(P_{41} < w, P_2 \leq 0, P_4 < 0.39)
\]

where:

\[
\begin{pmatrix}
P_{11} \\
P_2 \\
P_3
\end{pmatrix}
\sim N
\begin{pmatrix}
0, \\
\begin{pmatrix}
1 & 0 & -\lambda \\
0 & 1 & \frac{1}{2}
\end{pmatrix} \\
\begin{pmatrix}
-\lambda \\
\frac{1}{2}
\end{pmatrix}
\end{pmatrix}
\]
Let $w_{\alpha}$ be the upper $100\alpha$ percent quantile of $F_0$,

$$w_{\alpha} = F_{\alpha}^{-1}(1 - \alpha |H_0) = F_0^{-1}(1 - \alpha)$$

The type I error rate of the design can be controlled at level $\alpha$ if the 1-sided rejection region is $\Omega = \{W: W > w_{\alpha}\}$.

It can be seen that the critical rejection value $w_{\alpha}$ is a function of interim information time $\frac{n_1}{N}$ and $\rho$.

To calculate the critical rejection values $w_{\alpha}$ for the biomarker informed add-arm design with 4 active treatment arms, we develop an R function `addarm_cv_K5` (Appendix A).

This function can be used to calculate the critical rejection values for any values of $n_1, N, \rho$ and $\alpha$. Table 4.1 lists the values of $w_{0.025}$ for the biomarker informed add-arm design with different values of interim information time $\frac{n_1}{N}$ and $\rho$. As expected, the critical
rejection value $w_{0.025}$ increases as $\rho$ increases. Further, it can be seen that, when the interim information time increases, the critical rejection value $w_{0.025}$ increases.

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>0</th>
<th>0.2</th>
<th>0.5</th>
<th>0.8</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>$w_{0.025}$</td>
<td>1.96</td>
<td>2.02</td>
<td>2.101</td>
<td>2.172</td>
<td>2.214</td>
</tr>
<tr>
<td>$n_1/N = 0.3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$w_{0.025}$</td>
<td>1.96</td>
<td>2.037</td>
<td>2.137</td>
<td>2.221</td>
<td>2.267</td>
</tr>
<tr>
<td>$n_1/N = 0.5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$w_{0.025}$</td>
<td>1.96</td>
<td>2.056</td>
<td>2.176</td>
<td>2.27</td>
<td>2.317</td>
</tr>
<tr>
<td>$n_1/N = 0.8$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4.2.5 Power

The biomarker informed add-arm design we proposed uses the biomarker indication of the primary endpoint for interim decision(s), and the assumption of a unimodal dose-response curve to avoid testing every candidate dose level. Some poor performing doses can be identified, and hence patients need not be exposed to these doses.

In this section, we compare our proposed design against the biomarker informed two-stage winner design. The biomarker informed two-stage winner design was proposed and studied in Shun et al. (2008). It starts with several active treatment arms and a control arm. At interim, the inferior arms will be terminated based upon results for biomarker, and only the most promising treatment (“winner”) will be retained and carried to the end of the study with the control arm. The final comparison between the winner arm and the control arm will be performed on data from both stages and on study primary endpoint. In biomarker informed two-stage winner design, patients need to be exposed to every
candidate dose, though only a small number of patients will be exposed to the poor performing doses. We consider a variety of settings, including different kinds of correlations between the biomarker and the primary endpoint and different shapes of the unimodal dose-response curve, and compare the power performance of the biomarker informed add-arm design with the biomarker informed two-stage winner design. Our proposed design is shown to outperform the biomarker informed two-stage winner design on average.

4.2.5.1 Power-Performance Simulation: Design

As presented in Wang et al. (2013), the power of a biomarker informed adaptive clinical trial design depends on the shape of mean-level correlation between the biomarker and the primary endpoint, the uncertainty about the shape, and the individual-level correlation ρ between the two endpoints. It was shown that, a biomarker informed two-stage winner design is a better choice than a corresponding classical design when the shape of the mean-level correlation is as that is shown in Figure 4.3, and the uncertainty $\sigma^2 \leq 0.2$. Hence for the power-performance comparison between our proposed design and the biomarker informed two-stage winner design, we consider the two shapes of mean-level correlation in Figure 4.3, different values of $\sigma^2$ ($\sigma^2 \leq 0.2$) and different values of ρ.
The power performance of a biomarker informed two-stage winner design is not affected by the dose order. However, in the biomarker informed add-arm design we proposed, a different dose order implies a different dose-response relationship and a different testing order, hence the power of the design might be different. In the comparison of the power performance between our proposed design and the biomarker informed two-stage winner design, we therefore also consider different dose orders, that is, different shapes of the unimodal dose-response curve. The 3 different unimodal dose-response curves we considered are shown in Figure 4.4.
In our simulation, we assume $u^*_0 = 1, \sigma^2_{X} = 1, u^*_0 = 1, \sigma^2_{Y} = 1$ – this could always be achieved by scaling. Assume the standardized mean response of primary endpoint for the 4 active treatment groups are 1.1, 1.3, 1.4 and 1.5, respectively. If $u^*_X$ and $u^*_Y$ are linearly related, then the standardized mean response of biomarker for the 4 active treatment groups are 1.1, 1.3, 1.4 and 1.5, respectively. If $u^*_X$ and $u^*_Y$ are nonlinearly related as in shown Figure 4.3, assume the standardized mean response of biomarker for the 4 active treatment groups are 1.2, 1.7, 2 and 2.1, respectively.

Consider the biomarker informed designs with total sample size 420. This sample size will yield 85% power of a corresponding classical design with no interim adaptation using Dunnett’s test for multiplicity adjustment. Assume the interim adaptations based on biomarker response take place when the information time is 0.5 (i.e. $\frac{n_1}{N} = 0.5$). The simulated power results for the biomarker informed add-arm design and the biomarker informed two-stage winner design under different scenarios are presented in next section.
4.2.5.2 Power-Performance Simulation: Results

Table 4.2-Table 4.5 list the simulated power results for the two designs.

Table 4.2 Power: Biomarker Informed Two-Stage Winner Design (Linear)

<table>
<thead>
<tr>
<th>$\sigma^2_f$</th>
<th>$\rho$</th>
<th>0</th>
<th>0.2</th>
<th>0.5</th>
<th>0.8</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>91.4%</td>
<td>92.2%</td>
<td>92.3%</td>
<td>93.3%</td>
<td>93.8%</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>88.7%</td>
<td>88.7%</td>
<td>88.7%</td>
<td>90.2%</td>
<td>90.6%</td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>84.5%</td>
<td>84.7%</td>
<td>85.8%</td>
<td>86.5%</td>
<td>86.6%</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.3 Power: Biomarker Informed Two-Stage Winner Design (Nonlinear)

<table>
<thead>
<tr>
<th>$\sigma^2_f$</th>
<th>$\rho$</th>
<th>0</th>
<th>0.2</th>
<th>0.5</th>
<th>0.8</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>93.6%</td>
<td>93.7%</td>
<td>93.9%</td>
<td>94.2%</td>
<td>94.9%</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>88.6%</td>
<td>88.6%</td>
<td>88.8%</td>
<td>89.1%</td>
<td>89.8%</td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>82.3%</td>
<td>82.6%</td>
<td>82.6%</td>
<td>82.9%</td>
<td>83.9%</td>
<td></td>
</tr>
</tbody>
</table>
Table 4.4 Power: Biomarker Informed Add-Arm Design (Linear)

<table>
<thead>
<tr>
<th>Power</th>
<th>Shape 1</th>
<th>Shape 2</th>
<th>Shape 3</th>
<th>0</th>
<th>0.2</th>
<th>0.5</th>
<th>0.8</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\sigma^2_r$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>88.6%</td>
<td>92.6%</td>
<td>94.8%</td>
<td>85.7%</td>
<td>89.9%</td>
<td>93.0%</td>
<td>83.2%</td>
</tr>
<tr>
<td></td>
<td>0.2</td>
<td>89.6%</td>
<td>93.4%</td>
<td>95.1%</td>
<td>86.1%</td>
<td>89.9%</td>
<td>93.1%</td>
<td>83.4%</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>90.1%</td>
<td>93.9%</td>
<td>95.5%</td>
<td>87.3%</td>
<td>91.0%</td>
<td>93.4%</td>
<td>83.6%</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>90.8%</td>
<td>94.0%</td>
<td>95.8%</td>
<td>88.2%</td>
<td>92.2%</td>
<td>93.6%</td>
<td>85.1%</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>91.2%</td>
<td>95.2%</td>
<td>96.2%</td>
<td>88.5%</td>
<td>92.4%</td>
<td>94.0%</td>
<td>85.8%</td>
</tr>
</tbody>
</table>
Table 4.5 Power: Biomarker Informed Add-Arm Design (Nonlinear)

<table>
<thead>
<tr>
<th>Power</th>
<th>Shape 1</th>
<th>Shape 2</th>
<th>Shape 3</th>
<th>( \rho )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma^2_{\tau} )</td>
<td>0</td>
<td>0.2</td>
<td>0.5</td>
<td>0.8</td>
</tr>
<tr>
<td>0</td>
<td>96.3%</td>
<td>96.6%</td>
<td>96.6%</td>
<td>96.9%</td>
</tr>
<tr>
<td></td>
<td>95%</td>
<td>95.7%</td>
<td>95.9%</td>
<td>96%</td>
</tr>
<tr>
<td></td>
<td>96.5%</td>
<td>96.6%</td>
<td>96.6%</td>
<td>97.2%</td>
</tr>
<tr>
<td>0.1</td>
<td>93.1%</td>
<td>93.1%</td>
<td>93.1%</td>
<td>93.2%</td>
</tr>
<tr>
<td></td>
<td>90.8%</td>
<td>90.9%</td>
<td>91.2%</td>
<td>92%</td>
</tr>
<tr>
<td></td>
<td>92.5%</td>
<td>92.9%</td>
<td>93.1%</td>
<td>93.3%</td>
</tr>
<tr>
<td>0.2</td>
<td>88.4%</td>
<td>88.5%</td>
<td>88.3%</td>
<td>88.5%</td>
</tr>
<tr>
<td></td>
<td>85.5%</td>
<td>86.3%</td>
<td>86.4%</td>
<td>86.8%</td>
</tr>
<tr>
<td></td>
<td>87.6%</td>
<td>87.7%</td>
<td>87.9%</td>
<td>88.2%</td>
</tr>
</tbody>
</table>

Table 4.2 and Table 4.3 report the power of the biomarker informed two-stage winner design (4 active treatment arms and 1 control) for different values of \( \rho \) and \( \sigma^2_{\tau} \) when \( u^{x*} \) and \( u^{y*} \) are linearly and nonlinearly related respectively. The simulation results are consistent with what was found in Wang et al. (2013): when \( \sigma^2_{\tau} \) is fixed, the power mild increases as \( \rho \) increases; however, the power drops significantly as \( \sigma^2_{\tau} \) increases for a fixed \( \rho \). The biomarker informed two-stage winner design is in general better than a classical design when \( \sigma^2_{\tau} \leq 0.2 \) if \( u^{x*} \) and \( u^{y*} \) are linearly related; however, when \( u^{x*} \)
and $u^Y*$ are nonlinearly related as shown in Figure 4.3, the biomarker informed two-stage winner design will be a better choice only when $\sigma^2 \leq 0.1$.

Table 4.4 and Table 4.5 present the power of the corresponding biomarker informed add-arm design with different dose-response curves when $u^X*$ and $u^Y*$ are linearly and nonlinearly related respectively. As seen, $\rho$ only plays little role on power of the biomarker informed add-arm design, while $\sigma^2$ plays a significant role.

Further, we can see that the shape of dose-response curve and the shape of mean level correlation affect the power performance of the biomarker informed add-arm design.

When the mean level correlation between biomarker and primary endpoint is linear, i.e., $u^X*$ and $u^Y*$ are linearly related, the design with dose-response curve in shape 3 of Figure 4.4 has the highest power, while the design with dose-response curve in shape 1 has the worst power. Comparing Table 4.4 with Table 4.2, the power of a biomarker informed add-arm design with dose-response shape 3 is generally 2%~5% higher than a corresponding biomarker informed two-stage winner design; the power of the design with dose-response shape 2 is 1%~2% higher than the corresponding biomarker informed two-stage winner design; while the power of the design with dose-response shape 1 is 1%~3% lower than the corresponding biomarker informed two-stage winner design. When the mean level correlation between the biomarker and primary endpoint is nonlinear as shown in Figure 4.3, for all the 3 dose-response curves we considered, the biomarker informed add-arm design has better performance on power. According to our simulation, the biomarker informed add-arm design is 2%~5% higher on power than a corresponding biomarker informed two-stage winner design.
On average, the biomarker informed add-arm design we proposed is better than the corresponding biomarker informed two-stage winner design in terms of power.

### 4.3 Biomarker Informed Add-Arm Design with \( K (K > 4) \) Active Treatments

In this section, we describe how to carry out the biomarker informed add-arm design for a clinical trial with more than 4 active treatments. Designs with up to 7 active treatments are proposed.

#### 4.3.1 Biomarker Informed Add-Arm Design with 5 Active Treatments

Figure 4.5 shows the dose selection procedures for the biomarker informed add-arm design with 5 active treatment groups.

*Figure 4.5 Biomarker Informed Add-Arm Design for Unimodal Response (5+1 Case)*

Figure 4.5 shows the dose selection procedures for the biomarker informed add-arm design with 5 active treatment groups.
Stage 1: Assign $2.5n_1$ subjects in dose group 2, 0 and 4 using randomization ratio 1:0.5:1.

Stage 2: If the observations $x_{n_1}^{(2)} > x_{n_1}^{(4)}$, assign $2.5n_1$ subjects in dose group 1, 0 and 3 using 1:0.5:1 randomization. Otherwise, if the observations $x_{n_1}^{(2)} \leq x_{n_1}^{(4)}$, assign $2.5n_1$ subjects in dose group 3, 0 and 5 using 1:0.5:1 randomization.

Stage 3: (a) If $x_{n_1}^{(2)} > x_{n_1}^{(4)}$ and $\max(\frac{x_{n_1}^{(2)}}{x_{n_1}^{(3)}}, \frac{x_{n_1}^{(3)}}{x_{n_1}^{(4)}}) - \frac{x_{n_1}^{(1)}}{x_{n_1}^{(4)}} < c_R$, select dose 1 group as the best candidate dose group; otherwise, select dose 2 or dose 3 group as the best candidate dose group depending on which has a larger response. If $x_{n_1}^{(2)} \leq x_{n_1}^{(4)}$ and $\max(\frac{x_{n_1}^{(3)}}{x_{n_1}^{(4)}}, \frac{x_{n_1}^{(4)}}{x_{n_1}^{(5)}}) - \frac{x_{n_1}^{(5)}}{x_{n_1}^{(4)}} < c_R$, select dose 5 group as the best candidate dose group; otherwise, select dose 4 or dose 3 group as the best candidate depending on which has a larger response. (b) Assign $2n_2$ subjects in dose group 0 and the selected best candidate dose group using 1:1 randomization.

In total, there will be $5n_1 + 2n_2$ subjects. The final analysis will use measurements of study primary endpoint from $n_1 + n_2$ subjects in the selected best candidate dose group and the control group.

The default selection probabilities when $c_R = 0$ for the 5 doses in this design are:

$\frac{1}{8}, \frac{2}{8}, \frac{2}{8}, \frac{2}{8}, \frac{1}{8}$, respectively. To ensure every dose candidate has an equal selection probability when the response curve is flat (i.e., under $H_0$: $u_1^{Y^*} = u_2^{Y^*} = u_3^{Y^*} = u_4^{Y^*} = u_5^{Y^*}$), the following equation has to be hold:
Let $D_1 = \frac{X_n^{(2)} - X_n^{(4)}}{\sqrt{\frac{2\sigma_X^2}{n_1}}}, D_2 = \frac{X_n^{(2)} - X_n^{(1)}}{\sqrt{\frac{2\sigma_X^2}{n_1}}}, D_3 = \frac{X_n^{(3)} - X_n^{(1)}}{\sqrt{\frac{2\sigma_X^2}{n_1}}}$, the conditional probability can be written as:

$$P \left( X_n^{(2)} > X_n^{(4)}, \max \left( X_n^{(2)}, X_n^{(1)} \right) - X_n^{(1)} < c_R; H_0' \right)$$

$$= P \left( D_1 > 0, D_2 < \frac{c_R}{\sqrt{\frac{2\sigma_X^2}{n_1}}}, D_3 < \frac{c_R}{\sqrt{\frac{2\sigma_X^2}{n_1}}}; H_0' \right)$$

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{\frac{c_R}{\sqrt{\frac{2\sigma_X^2}{n_1}}}} \int_{-\frac{c_R}{\sqrt{\frac{2\sigma_X^2}{n_1}}} \text{ to } +\infty} f(d_1, d_2, d_3) d_d3 d_d2 d_d1$$

$$\text{where } \begin{pmatrix} D_1 \\ D_2 \\ D_3 \end{pmatrix} \sim N \left( \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \frac{1}{2} & 0 \\ \frac{1}{2} & 1 & \frac{1}{2} \\ 0 & \frac{1}{2} & 1 \end{pmatrix} \right) \text{ under } H_0'. f(d_1, d_2, d_3) \text{ is p.d.f. of } (D_1, D_2, D_3)'$$.

The integration could be carried out by the numerical software. It shows that, $\frac{c_R}{\sqrt{\frac{2\sigma_X^2}{n_1}}} = 0.37$. Thus, $c_R$ is a function of $n_1$ and $\sigma_X^2$, $c_R = 0.37 \sqrt{\frac{2\sigma_X^2}{n_1}}$.

$\sigma_X^2$ is suggested to be substituted by the sample variance $\overline{\sigma_X^2}$ for calculating $c_R$. The obtained $c_R$ is further suggested to be validated via simulations to ensure an (approximately) equal selection probability of each dose group for a flat response curve.
4.3.2 Biomarker Informed Add-Arm Design with 6 Active Treatments

The dose selection procedures for the biomarker informed add-arm design with 6 active treatment groups are shown in Figure 4.6.

Stage 1: Assign $2.5n_1$ subjects in dose group 3, 0 and 4 using randomization ratio 1:0.5:1.

Stage 2: If the observations \( \bar{x}^{(3)}_{n_1} > \bar{x}^{(4)}_{n_1} \), assign $2.5n_1$ subjects in dose group 1, 0 and 2 using 1:0.5:1 randomization. Otherwise, if the observations \( \bar{x}^{(3)}_{n_1} \leq \bar{x}^{(4)}_{n_1} \), assign $2.5n_1$ subjects in dose group 5, 0 and 6 using 1:0.5:1 randomization.

Stage 3: (a) If \( \bar{x}^{(3)}_{n_1} > \bar{x}^{(4)}_{n_1} \) and \( \bar{x}^{(3)}_{n_1} - \max(\bar{x}^{(1)}_{n_1},\bar{x}^{(2)}_{n_1}) > c_R \), select dose 3 group as the best candidate dose group; otherwise, select dose 1 or dose 2 group as the best candidate dose group depending on which has a larger response. If \( \bar{x}^{(3)}_{n_1} \leq \bar{x}^{(4)}_{n_1} \) and \( \bar{x}^{(4)}_{n_1} -
max \( \left( \frac{x_{n_1}^{(5)}}{c_R}, \frac{x_{n_1}^{(6)}}{c_R} \right) \) > 0, select dose 4 group as the best candidate dose group; otherwise, select dose 5 or dose 6 group as the best candidate depending on which has a larger response. (b) Assign \( 2n_2 \) subjects in dose group 0 and the selected best candidate dose group using 1:1 randomization.

In total, there will be \( 5n_1 + 2n_2 \) subjects. The final analysis will use measurements of study primary endpoint from \( n_1 + n_2 \) subjects in the selected best candidate dose group and the control group.

The default selection probabilities when \( c_R = 0 \) in this design are: \( \frac{1}{8}, \frac{1}{8}, \frac{2}{8}, \frac{2}{8}, \frac{1}{8}, \frac{1}{8} \), for the 6 doses, respectively. To make every dose candidate has an equal selection probability for a flat response curve (i.e., under the hypothesis \( H_0' \): \( u_4^{Y^*} = u_5^{Y^*} = u_6^{Y^*} = u_4^{Y^*} = u_5^{Y^*} = u_6^{Y^*} \)), the following equation need to be hold:

\[
P \left( \frac{x_{n_1}^{(3)}}{x_{n_1}^{(4)}}, \frac{x_{n_1}^{(3)}}{x_{n_1}^{(4)}} - \max \left( \frac{x_{n_1}^{(1)}}{x_{n_1}^{(2)}}, \frac{x_{n_1}^{(1)}}{x_{n_1}^{(2)}} \right) > c_R; H_0' \right) = \frac{1}{6}
\]

Let \( D_1 = \frac{x_{n_1}^{(3)} - x_{n_1}^{(4)}}{2\sigma_X^{\text{R}}/n_1}, D_2 = \frac{x_{n_1}^{(3)} - x_{n_1}^{(4)}}{2\sigma_X^{\text{R}}/n_1}, D_3 = \frac{x_{n_1}^{(3)} - x_{n_1}^{(2)}}{2\sigma_X^{\text{R}}/n_1} \), the conditional probability can be written as:

\[
P \left( \frac{x_{n_1}^{(3)}}{x_{n_1}^{(4)}}, \frac{x_{n_1}^{(3)}}{x_{n_1}^{(4)}} - \max \left( \frac{x_{n_1}^{(1)}}{x_{n_1}^{(2)}}, \frac{x_{n_1}^{(1)}}{x_{n_1}^{(2)}} \right) > c_R; H_0' \right)
\]

\[
= P \left( D_1 > 0, D_2 > \frac{c_R}{2\sigma_X^{\text{R}}/n_1}, D_3 > \frac{c_R}{2\sigma_X^{\text{R}}/n_1}; H_0' \right)
\]

\[
= \int_0^{+\infty} \int_0^{c_R/2\sigma_X^{\text{R}}/n_1} f(d_1, d_2, d_3) d_3 d_2 d_1
\]
where \( \begin{pmatrix} D_1 \\ D_2 \\ D_3 \end{pmatrix} \sim N \left( \begin{pmatrix} 0 \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}, \begin{pmatrix} 1 & 1/2 & 1/2 \\ 1/2 & 1 & 1/2 \\ 1/2 & 1/2 & 1 \end{pmatrix} \right) \) under \( H_0' \). \( f(d_1, d_2, d_3) \) is p.d.f. of \((D_1, D_2, D_3)'\).

By numerical software, \( \frac{c_R}{\sqrt{2\sigma^2_X/n_1}} = 0.355 \). Thus, \( c_R \) is a function of \( n_1 \) and \( \sigma^2_X \), \( c_R = 0.355 \frac{2\sigma^2_X}{\sqrt{n_1}} \).

\( \sigma^2_X \) is suggested to be substituted by the sample variance \( \overline{\sigma^2_X} \) for calculating \( c_R \). The obtained \( c_R \) is further suggested to be validated via simulations to ensure an (approximately) equal selection probability of each dose group for a flat response curve.

### 4.3.3 Biomarker Informed Add-Arm Design with 7 Active Treatments

The dose selection procedures for the biomarker informed add-arm design with 7 active treatment groups are shown in Figure 4.7.

![Figure 4.7 Biomarker Informed Add-Arm Design for Unimodal Response (7+1 Case)](image-url)
Stage 1: Assign 2.5\(n_1\) subjects in dose group 3, 0 and 5 using randomization ratio 1:0.5:1.

Stage 2: If the observations \(\bar{x}_{n_1}^{(3)} > \bar{x}_{n_1}^{(5)}\), assign 3.5\(n_1\) subjects in dose group 1, 2, 4 and 0 using 1:1:1:0.5 randomization. Otherwise, if the observations \(\bar{x}_{n_1}^{(3)} \leq \bar{x}_{n_1}^{(5)}\), assign 3.5\(n_1\) subjects in dose group 4, 6, 7 and 0 using 1:1:1:0.5 randomization.

Stage 3: (a) If \(\bar{x}_{n_1}^{(3)} > \bar{x}_{n_1}^{(5)}\) and \(\max(\bar{x}_{n_1}^{(1)}, \bar{x}_{n_1}^{(2)}) > \max(\bar{x}_{n_1}^{(3)}, \bar{x}_{n_1}^{(4)}) - c_R\), select dose 1 or dose 2 as the best candidate dose depending on which has a larger response; otherwise, select dose 3 or dose 4 as the best candidate dose depending on which has a larger response. If \(\bar{x}_{n_1}^{(3)} \leq \bar{x}_{n_1}^{(5)}\) and \(\max(\bar{x}_{n_1}^{(6)}, \bar{x}_{n_1}^{(7)}) > \max(\bar{x}_{n_1}^{(5)}, \bar{x}_{n_1}^{(4)}) - c_R\), select dose 7 or dose 6 group as the best candidate dose group; otherwise, select dose 5 or dose 4 group as the best candidate depending on which has a larger response. (b) Assign 2\(n_2\) subjects in dose group 0 and the selected best candidate dose group using 1:1 randomization.

In total, there will be \(6n_1 + 2n_2\) subjects required. The final analysis will compare measurements of study primary endpoint from \(n_1 + n_2\) subjects in the selected best candidate dose group and the control group.

The default selection probabilities when \(c_R = 0\) for the 7 doses in this design are:

\[
\begin{array}{cccccc}
\frac{1}{10} & \frac{1}{10} & \frac{2}{10} & \frac{2}{10} & \frac{1}{10} & \frac{1}{10} \\
\end{array}
\]

respectively. To ensure every dose candidate has an equal selection probability for a flat response curve (i.e., under the hypothesis \(H'_0\); \(u_1^{Y^*} = u_2^{Y^*} = u_3^{Y^*} = u_4^{Y^*} = u_5^{Y^*} = u_6^{Y^*} = u_7^{Y^*}\)), the following equation need to be hold:

\[
P(\bar{X}_{n_1}^{(3)} > \bar{X}_{n_1}^{(5)}, \max(\bar{X}_{n_1}^{(1)}, \bar{X}_{n_1}^{(2)}) > \max(\bar{X}_{n_1}^{(3)}, \bar{X}_{n_1}^{(4)}) - c_R, \bar{X}_{n_1}^{(1)} > \bar{X}_{n_1}^{(2)}, H'_0) = \frac{1}{7}
\]
Let $D_1 = \frac{X_{n_1}^{(3)} - X_{n_1}^{(5)}}{\sqrt{2\sigma_X^2/n_1}}$, $D_2 = \frac{X_{n_1}^{(1)} - X_{n_1}^{(3)}}{\sqrt{2\sigma_X^2/n_1}}$, $D_3 = \frac{X_{n_1}^{(1)} - X_{n_1}^{(4)}}{\sqrt{2\sigma_X^2/n_1}}$, $D_4 = \frac{X_{n_1}^{(1)} - X_{n_1}^{(2)}}{\sqrt{2\sigma_X^2/n_1}}$, the conditional probability can be written as:

$$P\left( X_{n_1}^{(3)} > X_{n_1}^{(5)}, \max\left(X_{n_1}^{(1)}, X_{n_1}^{(2)}\right) > \max\left(X_{n_1}^{(3)}, X_{n_1}^{(4)}\right) - c_R, X_{n_1}^{(1)} > X_{n_1}^{(2)}; H_0^\prime \right)$$

$$= P\left( D_1 > 0, D_2 > -c_R, D_3 > -c_R, D_4 > 0; H_0^\prime \right)$$

$$= \int_0^{+\infty} \int_{-\infty}^{c_R^2} \int_0^{c_R^2} \int_{-\infty}^{+\infty} f(d_1, d_2, d_3, d_4) \, d_4 \, d_3 \, d_2 \, d_1$$

where

$$\begin{pmatrix} D_1 \\ D_2 \\ D_3 \\ D_4 \end{pmatrix} \sim N\left( \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & -1/2 & 0 & 0 \\ -1/2 & 1 & 1/2 & 1/2 \\ 0 & 1/2 & 1 & 1/2 \\ 0 & 1/2 & 1/2 & 1 \end{pmatrix} \right)$$

under $H_0^\prime$. $f(d_1, d_2, d_3, d_4)$ is p.d.f. of $(D_1, D_2, D_3, D_4)^\prime$.

By numerical software, $c_R = 0.339$. Thus, $c_R$ is a function of $n_1$ and $\sigma_X^2$, $c_R = 0.339 \sqrt{2\sigma_X^2/n_1}$.

$\sigma_X^2$ is suggested to be substituted by the sample variance $\overline{\sigma_X^2}$ for calculating $c_R$. The obtained $c_R$ is further suggested to be validated via simulations to ensure an (approximately) equal selection probability of each dose group for a flat response curve.
In addition to comparing the observed means of biomarker at the interim stages for dose selection, comparing the observed values of t-statistic might be another option for interim decisions of a biomarker informed add-arm design. In a biomarker informed add-arm design that uses the observed t-statistic for interim decisions, $c_R$ would be a constant that does not vary with $n_1$ and $\sigma^2_X$.

4.4 Discussion

In this chapter, we have proposed a new biomarker informed adaptive design called “biomarker informed add-arm design for unimodal response”. Designs starting with 4 active treatment arms up to 7 active treatment arms were considered. This design combines the initial dose selection based on a biomarker and confirmation based on the primary endpoint into the same trial. It utilizes the existing knowledge about the shape of dose-response of the biomarker and determines the test order for doses to select the most efficacious dose for the next stage. It is different from the biomarker informed two-stage winner design where all treatments need to be explored while the biomarker add-arm design can identify some inferior treatments and hence no patients need to be exposed to those doses. Consequently, the proposed design may further reduce the expected sample size while gaining information about the effective dose more quickly.

We discussed the approach for ensuring each dose has equal selection probability under flat response curve and proposed a statistical approach for controlling type I error of the
design. R function for calculating the adjusted critical rejection value for final test statistic of the design was developed.

We carried out extensive simulation studies to study this design and to evaluate its performance against the corresponding biomarker informed two-stage winner design. It has been shown that the power of our proposed design is mainly affected by the shape of dose-response curve and the shape of mean-level correlation and its uncertainty. The power of the design is in general better, when the mode of the dose-response curve is in the middle, and the shape of mean-level correlation is nonlinear as shown in Figure 4.3. On average, the proposed design is better than its corresponding biomarker informed two-stage winner design. However, the benefits are different for different shapes of dose-response, and the mean-level relationship.

Our proposed design has the potential to efficiently use the existing information and resources and to select the most effective dose faster, but it suffers with the same limitation as any other biomarker informed adaptive designs, as it requires certain certainty about the surrogacy of the biomarker in order to ensure the advantage of the design. Though we have incorporated the parameters of uncertainty by using the two-level correlation model proposed by Wang, et al. (2013) for design evaluation, careful consideration is strongly suggested before employing the proposed design.
CHAPTER 5

A CASE STUDY OF USING BIOMARKER AT INTERIM

5.1 Introduction

Biomarker informed adaptive clinical trial designs, which make interim decisions based upon inference on biomarkers, have generated a great deal of attention in recent years. Using biomarkers for interim decisions can help forestall certain operational challenges, and have the potential to shorten the duration of the trial for drug development and prevent unnecessary waste of resources. But how much would we pay for this?

In this chapter, we perform simulation studies for non-small cell lung cancer trials where change in tumor size is used as a biomarker of survival. The concordance rate of decisions made based upon survival times and changes in tumor size is investigated, and the factors that affect the concordance of decisions are explored. This study will help us better understand the profit and loss of using a biomarker for interim decisions.

Different from previous chapters where we specify distributions for both endpoints in the study, a regression model linking the biomarker and primary endpoint will be specified to describe the relationship of the two endpoints.

5.2 Survival Model in Non-Small-Cell Lung Cancer

Change in tumor size can be used as a biomarker for survival in non-small-cell lung cancer.
Wang et al. (2009) developed a parametric model for survival times in non-small-cell lung cancer patients, utilizing data from four non-small-cell lung cancer registration trials. The parametric survival model they proposed includes baseline tumor size (centered at 8.5 cm), ECOG status (0/1/2/3 as a categorical variable) and percentage tumor reduction from baseline at week 8 ($PTR_{wk8}$) as predictors of time to death ($T$). The regression model is written as follows:

$$\log(T) = \alpha_0 + \alpha_1 \times ECOG + \alpha_2 \times (Baseline - 8.5) + \alpha_3 \times PTR_{wk8} + \varepsilon_{TD}$$

where $T$ is the time to death (day), $\alpha_0$ is the intercept, $\alpha_1$, $\alpha_2$, $\alpha_3$ are the slopes for ECOG, centered baseline, and $PTR_{wk8}$, respectively, and $\varepsilon_{TD}$ is the residual variability following a normal distribution with a mean of 0 and variance of $\sigma_{TD}^2$.

A set of parameter estimates has been provided for the second-line treatments of non-small-cell lung cancer:

$\hat{\alpha}_0 = 5.9$

$\hat{\alpha}_1 = -0.3$

$\hat{\alpha}_2 = -0.03$

$\hat{\alpha}_3 = 0.4$

Thus,

$$\log(T) = 5.9 - 0.3 \times ECOG - 0.03 \times (Baseline - 8.5) + 0.4 \times PTR_{wk8} + \varepsilon_{TD} \quad (5.1)$$

This model has been shown to have good predictive ability by Wang et al. (2009).

In the following, we utilize this model and carry out simulation studies to investigate the concordance between decisions made based upon $PTR_{wk8}$ and based on survival $T$ for a non-small-cell lung cancer trial using classical design.
5.3 Data Generation Based On Survival Model for Non-Small-Cell Lung Cancer

As presented in Chapter 3, the individual-level correlation, the estimated mean-level correlation, and the uncertainty about the estimated mean-level correlation are three factors that determine the relationship of a biomarker and a primary endpoint. When the model (5.1) is specified, the estimated mean-level correlation is fixed. For simplicity, we consider the most ideal scenario and further assume that there is no uncertainty about this estimated mean-level correlation. Therefore, the individual-level correlation \( \rho \) is the only index in our case for describing the relationship between the two endpoints.

A natural question is how to generate data for \( PTR_{wkB} \) and \( \log(T) \) so that they satisfy (5.1) and correlate with the correlation coefficient \( \rho \). We address this question in this section.

Assume \( \log(T) \sim N(u_T, \sigma_T^2) \), \( PTR_{wkB} \sim N(u_{PTR}, \sigma_{PTR}^2) \), \( \varepsilon_{TD} \sim N(0, \sigma_{TD}^2) \) and assume the mean ECOG score of patients is 0.5 and the baseline tumor size follows a normal distribution with mean 8.5 and variance 30, to make \( PTR_{wkB} \) and \( \log(T) \) correlate with a correlation coefficient \( \rho \) while preserving the above relationship (5.1), the following equations need to be satisfied (see proof in Appendix C):

\[
\begin{align*}
  u_T &= 5.75 + 0.4u_{PTR} \\
  \sigma_{PTR} &= \rho \frac{\sigma_T}{0.4} \\
  \sigma_{TD}^2 &= (1 - \rho^2)\sigma_T^2
\end{align*}
\]
By convention, we make assumptions for the primary study endpoint $\log(T)$ (i.e. $u_T, \sigma_T^2$). Data for $PTR_{wk8}$ and $\log(T)$ that satisfy (5.1) and correlate with a correlation coefficient $\rho$ could be generated by the following procedure:

1. calculate $u_{PTR}, \sigma_{PTR}^2,$ and $\sigma_{TD}^2$ using (5.2)-(5.4)
2. simulate $PTR_{wk8}$ from $N(u_{PTR}, \sigma_{PTR}^2),$ and $\varepsilon_{TD}$ from $N(0, \sigma_{TD}^2)$
3. calculate $\log(T)$ using (5.1)

For example, if we assume $u_T = \log(180)$ (i.e. the mean survival time is around 6 months), $\sigma_T^2 = 4,$ and $\rho=0.5,$ using the above procedure if we generate 500 pairs of data; the sample correlation coefficient is 0.505.

It is not recommended to simulate data for survival ($\log(T)$ ) and $\varepsilon_{TD}$ first, and calculate $PTR_{wk8}$ from (5.1). If data were generated in this way, the sample correlation coefficient for the above example would be 0.75. The correlation between $\log(T)$ and $PTR_{wk8}$ would not be preserved.

Follow-up times could also be generated from another log-normal distribution to control the censoring rate for survival.

5.4 Inference Based Upon Change in Tumor Size vs. Inference Based On Survival

In this section, we investigate the concordance between decisions made by the tumor size changes at week 8 ($PTR_{wk8}$) and by survival times ($\log(T)$ ) in a non-small-cell lung cancer trial. The factors that affect the concordance of decisions are also explored.
For a more accurate result, we assume the trial uses a classical design. That is, the same number of patients will be used for inference based on tumor size change and on survival. The “interim” here refers to the short time needed for the measurements of tumor size change at week 8.

Assume there are two arms, a treatment arm and a control arm. The patient populations in the two arms are assumed to share the same baseline characteristics, such as the baseline tumor size and the ECOG score. Assume the expected mean survival time for patients in the treatment arm is around 9 months (i.e., \( u_{T1} = \log(270) \)), and the expected mean survival time for patients in the control arm is around 6 months (i.e., \( u_{T0} = \log(180) \)).

Assume \( \sigma^2 = 2 \), then 191 patients are required in each arm to achieve 80% power of the design with 2-sided type I error rate controlled at 0.05.

We generate patients’ records for week 8 tumor size change \( PTR_{wk8} \) and survival \( T \) using the procedure stated in the previous section. Follow-up times are also generated from log-normal distributions to control the censoring rate. Four censoring rates we studied are 0%, 5%, 10%, and 20%. We use t-test for comparing week 8 tumor size changes \( PTR_{wk8} \) between the treatment arm and the control arm. For comparison of the survival times between the two arms, we consider two survival analysis methods: the log-rank test, which we use most of the time; and the survival analysis with lognormal distribution, when the survival distribution is specified correctly.

We compare the decisions made based upon t-test for \( PTR_{wk8} \), log-rank test for \( T \), and survival analysis with lognormal distribution for \( T \), and calculate the concordance rates of
decisions. The simulated results under different scenarios are listed in Table 5.1—Table 5.3.

Table 5.1 Concordance Rate: Design Specifications Assumed Correctly

<table>
<thead>
<tr>
<th>Concordance Rate</th>
<th>N=191, $\sigma^2_f = 2$, $u_{T0} = \log(180)$, $u_{T1} = \log(270)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>t vs. LR</td>
</tr>
<tr>
<td>$\rho = 0.9$</td>
<td></td>
</tr>
<tr>
<td>$\rho = 0.75$</td>
<td>0.82</td>
</tr>
<tr>
<td>$\rho = 0.5$</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>0.72</td>
</tr>
<tr>
<td>Censoring Rate</td>
<td></td>
</tr>
<tr>
<td>0%</td>
<td>0.80</td>
</tr>
<tr>
<td>5%</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>0.71</td>
</tr>
<tr>
<td>10%</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>0.72</td>
</tr>
<tr>
<td>20%</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>0.71</td>
</tr>
<tr>
<td>Concordance Rate</td>
<td>$\rho = 0.9$</td>
</tr>
<tr>
<td>------------------</td>
<td>--------------</td>
</tr>
<tr>
<td></td>
<td>N=191, $\sigma_T^2 = 4$, $u_{T_0} = \log(180)$, $u_{T_1} = \log(270)$</td>
</tr>
<tr>
<td></td>
<td>t vs. LR</td>
</tr>
<tr>
<td>Censoring Rate</td>
<td></td>
</tr>
<tr>
<td>0%</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>0.46</td>
</tr>
<tr>
<td>5%</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>0.46</td>
</tr>
<tr>
<td>10%</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>0.46</td>
</tr>
<tr>
<td>20%</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>0.63</td>
</tr>
<tr>
<td></td>
<td>0.45</td>
</tr>
<tr>
<td>Censoring Rate</td>
<td>Concordance Rate</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------</td>
</tr>
<tr>
<td></td>
<td>$\rho = 0.9$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>0%</td>
<td>$t$ vs. LR</td>
</tr>
<tr>
<td></td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>0.46</td>
</tr>
<tr>
<td>5%</td>
<td>$t$ vs. LR</td>
</tr>
<tr>
<td></td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>0.45</td>
</tr>
<tr>
<td>10%</td>
<td>$t$ vs. LR</td>
</tr>
<tr>
<td></td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>0.63</td>
</tr>
<tr>
<td></td>
<td>0.44</td>
</tr>
<tr>
<td>20%</td>
<td>$t$ vs. LR</td>
</tr>
<tr>
<td></td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>0.43</td>
</tr>
</tbody>
</table>
Table 5.1 lists the simulated concordance rates for different values of $\rho$ and censoring rate when the design specifications, including $\sigma^2_T$, $u_{T0}$, and $u_{T1}$, are assumed correctly. Table 5.2 lists the simulated concordance rates when $\sigma^2_T$ is assumed incorrectly: the true value for $\sigma^2_T$ is 4, but we mistakenly assume $\sigma^2_T = 2$. Table 5.3 lists the simulated concordance rates when the treatment effect is not assumed correctly: we assume a 3 months difference in survival between the two different treatments, when in fact there is only 2 months’ difference in survival.

We can see that, in general, censoring does not affect the concordance of decisions much if the censoring rate is controlled within 20%. By our simulations, the difference in concordance rates is within 3% for different values of the censoring rate.

In all the scenarios we considered, the concordance between decisions by the two different survival analysis techniques are around 85%~90%. Thus, a 10%~15% difference in inference decisions will arise if we perform survival analysis using a non-parametric method instead of specifying a correct survival distribution.

In general, the concordance between decisions made based on $PTR_{wk8}$ and decisions made based on $T$ using survival analysis with a correct survival distribution is higher than the concordance between decisions by $PTR_{wk8}$ and by $T$ using log-rank test.

Under our simulation settings, if all the design specifications ($\sigma^2_T$, $u_{T0}$, $u_{T1}$) are specified correctly, for a trial designed with 80% power, concordance of decisions by t-test for $PTR_{wk8}$ and by log-rank test for $T$ is also around 80% if the correlation between $PTR_{wk8}$ and $T$ equals 0.9, i.e., $\rho = 0.9$. When the correlation coefficient $\rho$ between the two
endpoints drops to 0.5, the concordance between decisions by $PTR_{wkB}$ and that by $T$ drop around 10%.

If any of the design specifications is specified incorrectly, the concordance between decisions made based on $PTR_{wkB}$ and based on $T$ will decrease. Moreover in this case, when $\rho$ drops from 0.9 to 0.5, the concordance rates of decisions drop around 30%. This indicates that sample size plays a role for the concordance between inference decisions based on interim biomarker and based on study primary endpoint when the mean-level relationship between the two endpoints is fixed.

If we were able to assume all the design specifications correctly, that is if we have a sufficient sample size, even if $\rho$ is small, the concordance rates are not terribly inaccurate in our case. For example in Table 5.1, when $\rho = 0.5$, the concordance between decisions based on $PTR_{wkB}$ and based on $T$ using log-rank test is 72%. However, if any of the design specifications is assumed incorrectly, the concordance rates will be very poor when $\rho$ is small.

5.5 Discussion

In this chapter, we studied the concordance between decisions made based on interim biomarker and based on study primary endpoint to better understand the profit and loss of using a biomarker for interim decisions.

In the context of a classical non-small-cell lung cancer trial where a regression model linking the interim biomarker (change in tumor size) and primary endpoint (survival) has been developed, we proposed a procedure for generating data for the two endpoints and
carried out extensive simulation studies. The inference decisions made based on tumor size changes and based on survival times using two different survival analysis techniques were compared under different scenarios, and the concordance rates were calculated. It has been shown that censoring of survival times does not affect the concordance of decisions much if the censoring rate is controlled within 20%. However, performing survival analysis using different methods will produce a mild difference in the inference decisions. Around 10%~15% difference in decisions will arise if we perform survival analysis using non-parametric method instead of specifying a correct survival distribution.

Sample size also plays a role in the concordance between inference decisions based on tumor size change and based on survival. When mean-level relationship between the two endpoints is fixed, if all the design specifications were assumed correctly (that is, if we have a sufficient sample size), the concordance of decisions would not terribly inaccurate even if the two endpoints correlated poorly. However, if any of the design specifications was specified incorrectly (meaning, if we underestimate the sample size required), the inference decisions based on a poor biomarker would be very different from that based on the primary endpoint. These findings agree with what we found in previous chapters.

Simulations in this chapter are based on the assumption that there is no uncertainty about the regression model linking the interim biomarker (change in tumor size) and primary endpoint (survival). Concordance of inference decisions might be worse if we are unsure about this relationship. Therefore, serious consideration is strongly suggested before using a biomarker for interim decisions.
CHAPTER 6

SUMMARY AND USING A BIOMARKER INFORMED ADAPTIVE DESIGN IN AN A&WC STUDY

In this dissertation, we develop novel statistical methods for biomarker informed adaptive clinical trial designs.

In Chapter 2, we propose a novel statistical approach for type I error control of the biomarker informed two-stage winner design. Our proposed method can work for the design with any number of treatment arms, which overcomes the limitation of the normal approximation method proposed by Shun et al. (2008). In Chapter 3, we develop a new two-level correlation approach to model the two endpoints in a biomarker informed adaptive design. A new variable that describes the mean level correlation is incorporated in the model. Compared with the conventional model which only considers the individual level correlation, the new model could reflect the uncertainty of the historical knowledge more accurately. In Chapter 4, we propose a new biomarker informed adaptive design called “biomarker informed add-arm design for unimodal response”. This design combines the initial dose selection based on a biomarker and confirmation based on the primary endpoint into the same trial. It utilizes the existing knowledge about the shape of dose-response curve and determines the test order for doses to select the most efficacious dose for the next stage. This design is shown to outperform the corresponding biomarker informed two-stage winner design in power. In Chapter 5, we study the concordance
between inference decisions made based on interim biomarker and based on study primary endpoint in the context of a classical non-small-cell lung cancer trial where a regression model linking the interim biomarker (change in tumor size) and primary endpoint (survival) exists. It is shown that sample size affects the concordance of inference decisions.

Biomarker informed adaptive clinical trial designs can help prevent certain operational challenges, and have the potential to shorten the duration of the trial for drug development and avoid unnecessary waste of resources. However, in order to ensure the benefits, certain certainty about the surrogacy of the biomarker is required. Therefore, a careful consideration is strongly suggested before employing a biomarker informed adaptive design. When a good portion of the relationship between the biomarker and the primary endpoint is known, the biomarker informed adaptive design is recommended.

In FDA Guidance for Industry about Adaptive Design Clinical Trials for Drugs and Biologics,

- Adaptation of study eligibility criteria based on analyses of pretreatment data
- Adaptations to maintain study power based on blinded interim analyses of aggregate data
- Adaptations based on interim results of an outcome unrelated to efficacy
- Adaptations using group sequential methods and unblinded analyses for early study termination because of either lack of benefit or demonstrated efficacy
- Adaptations in the data analysis plan not dependent on within study, between-group outcome differences are categorized as well-understood adaptive approaches, and can be used in adequate and well-controlled (A&WC) studies.

- Adaptations for dose selection studies
- Adaptive randomization based on relative treatment group responses
- Adaptations of sample size based on interim effect size estimates
- Adaptations of patient population based on treatment-effect estimates
- Adaptation for endpoint selection based on interim estimate of treatment effect
- Adaptation of multiple-study design features in a single study
- Adaptations in non-inferiority studies

are categorized as less well-understood adaptive approaches, and are encouraged to be applied in the exploratory studies to gain more experience.

It can be seen that, neither of the un-blinded adaptive designs (except the group sequential design) is categorized as the well-understood adaptive approach. The 3 major concerns by FDA of using an adaptive design in an adequate and well-controlled study include: type I error inflation, estimation bias, and operational bias.

For most of the less well-understood conventional adaptive designs listed above, the mathematical framework has been well established. Different statistical methods have been proposed in literature to prevent type I error inflation and estimation bias. The concerns for these designs are mainly for the operational challenges.
However, for the biomarker informed adaptive designs, even the mathematical framework has not been fully established yet. Then, is it possible to use a biomarker informed adaptive approach in an A&WC study? At the end of this dissertation, we describe a scenario where we think the biomarker informed adaptive approach might be used in the A&WC study.

The new trend of making cancer drugs is to develop the combination drugs, that contain a marketed drug (denoted by “A”) and an investigational novel drug (denoted by “P”). To get a combination drug (denoted by “A+P”) approved, besides showing the advantage of the combination drug A+P over the standard treatment (denoted by “T”) in market, it’s also required to show that A+P is better than A in order to demonstrate the added effect of novel drug A. It might take too long for the trial to start comparing A+P with T after seeing the added effect of P in a phase III trial. The biomarker informed adaptive approach might be utilized in this scenario for an early decision about when to start recruiting patients for the standard treatment arm T.

More specifically, start the trial with recruiting patients for the combination drug A+P and the marketed drug A, plan an interim look for comparison of the two arms (A+P vs. A) on biomarker (e.g. tumor size change). If promising results show on the biomarker, start recruiting patients for the standard treatment arm T; otherwise, start recruiting patients for the standard treatment arm T until the end of the study when the added effect of P is shown.

In this scenario, the adaptation is used for optimizing the operational procedures, which might not undermine the validity and integrity of the clinical trial.
APPENDIX A

R Functions for Calculation of Critical Rejection Values for
Biomarker Informed Two-Stage Winner Design and
Biomarker Informed Add-Arm Design for Unimodal Response

library(mnormt)

convention_cv_K4 <- function(w, n1, N, rho, alpha){
  K=4
  r=rho*sqrt(n1/(2*N))
  mu=rep(0, K-1)
  varcov=matrix(c(1,r,r,2,1,r,1,2),3,3)
  p=(K-1)*sadmvn(lower=c(-18, 0, 0), upper=c(w, 18, 18), mu, varcov)
  error=p[1]-(1-alpha)
  return(list(p=p, error=error))
}

convention_cv_K5 <- function(w, n1, N, rho, alpha){
  K=5
  r=rho*sqrt(n1/(2*N))
  mu=rep(0, K-1)
  varcov=matrix(c(1,rep(r,4),2,1,1,1,1,2,1, r, 1,1,2,4,4)
  p=(K-1)*sadmvn(lower=c(-18, 0, 0,0), upper=c(w, 18, 18,18), mu, varcov)
  error=p[1]-(1-alpha)
  return(list(p=p, error=error))
}

convention_cv_K6 <- function(w, n1, N, rho, alpha){
  K=6
  r=rho*sqrt(n1/(2*N))
  mu=rep(0, K-1)
  varcov=matrix(c(1,rep(r,5),2,1,1,1,1,2,1,1,1,1,2,1,1,1,1,2,5,5)
  p=(K-1)*sadmvn(lower=c(-18, 0, 0,0,0), upper=c(w, 18, 18,18,18), mu, varcov)
  error=p[1]-(1-alpha)
  return(list(p=p, error=error))
}

convention_cv_K7 <- function(w, n1, N, rho, alpha){
  K=7
  r=rho*sqrt(n1/(2*N))
mu=rep(0, K-1)
varcov=matrix(c(1,rep(r,6),2,rep(1,4),r,1,2,rep(1,3),r,1,1,2,1,1,1, rep(1,3),2,1,1, rep(1,4),2),6,6)
p=(K-1)*sadmvn(lower=c(-18, 0, 0,0,0,0), upper=c(w, 18,18,18,18,18), mu, varcov)
error=p[1]-(1-alpha)
return(list(p=p, error=error))
}

convention_cv_K8 <- function(w, n1, N, rho, alpha){
K=8
r=rho*sqrt(n1/(2*N))
mu=rep(0, K-1)
varcov=matrix(c(1,rep(r,7),2,rep(1,5),r,1,2,rep(1,4),r,1,1,2,1,1,1, rep(1,3),2,1,1, r, rep(1,5),2),7,7)
p=(K-1)*sadmvn(lower=c(-18, 0, 0,0,0,0,0), upper=c(w, 18,18,18,18,18,18,18), mu, varcov)
error=p[1]-(1-alpha)
return(list(p=p, error=error))
}

wang_cv_K4 <- function(w, n1,N, rho, u0, s2_0,s2_1,s2_2,s2_3,alpha){
K=4
seq_p=c()
mu=rep(0,K-1)
s2=c(s2_1,s2_2,s2_3)
for (j in 1:length(s2)){
a=s2[j]
r=rho*sqrt(n1/N)*sqrt((u0^2*a+1)/(u0^2*a+u0^2*s2_0+2))
varcov=matrix(c(1,r,r,r,2,1,r,1,2),3,3)
p=sadmvn(lower=c(-18, 0, 0), upper=c(w, 18), mu, varcov)
seq_p=c(seq_p,p[1])
}total_p=sum(seq_p)
error=total_p-(1-alpha)
return(list(p=total_p, error=error))
}

wang_cv_K5 <- function(w, n1,N, rho, u0, s2_0,s2_1,s2_2,s2_3,s2_4,alpha){
K=5
seq_p=c()
mu=rep(0,K-1)
s2=c(s2_1,s2_2,s2_3,s2_4)
for (j in 1:length(s2)){
a=s2[j]
\[ r = \rho \sqrt{\frac{n_1}{N}} \sqrt{\frac{(u_0^2 a + 1)}{(u_0^2 a + u_0^2 s_{2_0} + 2)}} \]

\[ \text{varcov} = \text{matrix}(c(1, \text{rep}(r, 4), 2, 1, 1, r, 1, 1, 2, 1, r, 1, 1, 2, 1, 2, 4, 4)) \]

\[ p = \text{sadmvn}(\text{lower} = c(-18, 0, 0, 0), \text{upper} = c(w, 18, 18, 18), \mu, \text{varcov}) \]

\[ \text{seq}_p = c(\text{seq}_p, p[1]) \]

\[ \text{total}_p = \text{sum}(\text{seq}_p) \]

\[ \text{error} = \text{total}_p - (1 - \alpha) \]

\[ \text{return}(\text{list}(p = \text{total}_p, \text{error} = \text{error})) \]

\textbf{wang_cv_K6} <- function(w, n1, N, rho, u0, s2_0, s2_1, s2_2, s2_3, s2_4, s2_5, alpha){
  K=6
  seq_p=c()
  mu=rep(0,K-1)
  s2=c(s2_1, s2_2, s2_3, s2_4, s2_5)
  for (j in 1:length(s2)){
    a=s2[j]
    r=rho*sqrt(n1/N)*sqrt((u0^2*a+1)/(u0^2*a+u0^2*s2_0+2))
    varcov=matrix(c(1, \text{rep}(r, 5), 2, 1, 1, r, 1, 1, 2, 1, r, 1, 1, 2, 1, 2, 5, 5))
    p=\text{sadmvn}(\text{lower} = c(-18, 0, 0, 0, 0), \text{upper} = c(w, 18, 18, 18, 18), \mu, \text{varcov})
    seq_p=c(seq_p, p[1])
  }
  total_p=\text{sum}(seq_p)
  error=total_p - (1 - \alpha)
  \text{return}(\text{list}(p = total_p, error = error))
}

\textbf{wang_cv_K7} <- function(w, n1, N, rho, u0, s2_0, s2_1, s2_2, s2_3, s2_4, s2_5, s2_6, alpha){
  K=7
  seq_p=c()
  mu=rep(0,K-1)
  s2=c(s2_1, s2_2, s2_3, s2_4, s2_5, s2_6)
  for (j in 1:length(s2)){
    a=s2[j]
    r=rho*sqrt(n1/N)*sqrt((u0^2*a+1)/(u0^2*a+u0^2*s2_0+2))
    varcov=matrix(c(1, \text{rep}(r, 6), 2, \text{rep}(1, 4), r, 1, 2, \text{rep}(1, 3), r, 1, 1, 2, 1, r, \text{rep}(1, 3), 2, 1, r, \text{rep}(1, 4), 2, 6, 6))
    p=\text{sadmvn}(\text{lower} = c(-18, 0, 0, 0, 0, 0), \text{upper} = c(w, 18, 18, 18, 18, 18), \mu, \text{varcov})
    seq_p=c(seq_p, p[1])
  }
  total_p=\text{sum}(seq_p)
  error=total_p - (1 - \alpha)
  \text{return}(\text{list}(p = total_p, error = error))
}
**wang_cv_K8** <- function(w, n1,N, rho, u0,
s2_0,s2_1,s2_2,s2_3,s2_4,s2_5,s2_6,s2_7,alpha){
  K=8
  seq_p=c()
  mu=rep(0,K-1)
  s2=c(s2_1,s2_2,s2_3,s2_4,s2_5,s2_6,s2_7)
  for (j in 1:length(s2)){
    a=s2[j]
    r=rho*sqrt(n1/N)*sqrt((u0^2*a+1)/(u0^2*a+u0^2*s2_0+2))
    varcov=matrix(c(1,rep(r,7),2,rep(1,5),r,1,1,2,rep(1,3),r, rep(1,3),2,1,1,r, rep(1,4),2,1, r, rep(1,5),2,7,7)
    p=sadmvn(lower=c(-18, 0, 0, 0,0,0, 0), upper=c(w, 18, 18, 18,18,18, 18), mu, varcov)
    seq_p=c(seq_p,p[1])
  }
  total_p=sum(seq_p)
  error=total_p
  return(list(p=total_p, error=error))
}

**addarm_cv_K5** <- function(w, n1=50, N=100, rho, alpha=0.025){
  K=5
  r=sqrt(n1/N)*rho/2
  mu=rep(0, K-2)
  varcov1=matrix(c(1,0,-r,0, 1, 0.5, -r, 0.5, 1),3,3)
  varcov2=matrix(c(1,r, r, 1, 0.5, r, 0.5,1),3,3)
  varcov3=matrix(c(1, -r, -r, -r, 1, 0.5, r, -0.5, 1),3,3)
  varcov4=matrix(c(1,-r,0, 1, -0.5, -r, 0.5, 1),3,3)
  p=sadmvn(lower=c(-18, 0, -18), upper=c(w, 18, 0.39), mu, varcov1)+sadmvn(lower=c(-18, 0, 0.39), upper=c(w, 18, 18), mu,
  varcov2)+sadmvn(lower=c(-18, -18, -18), upper=c(w, 0, 18), mu,
  varcov3)+sadmvn(lower=c(-18, -18, 0.39), upper=c(w, 0, 0.39),
  mu, varcov4)
  error=p[1]-(1-alpha)
  return(list(p=p, error=error))
}
APPENDIX B

Derivation for $\hat{\sigma}_{rj}^2$

$$\sigma_{rj}^2 = \text{Var} \left( \frac{\bar{Y}(j)/\sigma_Y}{\bar{X}(j)/\sigma_X} \right)$$

$$\approx \frac{\sigma_X^2}{\sigma_Y^2} \left( \frac{E^2 \bar{Y}(j)}{E^2 \bar{X}(j)} \left[ \text{Var} \bar{Y}(j) - 2 \frac{\text{Cov}(\bar{X}(j) \bar{Y}(j))}{E \bar{X}(j) E \bar{Y}(j)} + \text{Var} \bar{X}(j) \right] \right)$$

$$= \frac{\sigma_X^2}{\sigma_Y^2} \left( \left( \frac{u_Y^r}{u_X^r} \right)^2 + \frac{\sigma_X^2}{n_j} \left[ \frac{\sigma_X^2}{n_j} - 2 \frac{1}{n_j} \rho \sigma_X \sigma_Y \left( \frac{1}{u_j^r} \right) - \frac{1}{n_j} \right] \right)$$

$$= \left( \frac{u_Y^r}{u_X^r} \right)^2 + \frac{1}{n_j} \left[ \frac{1}{n_j} - 2 \frac{1}{n_j} \rho \sigma_X \sigma_Y + \frac{1}{n_j} \right]$$

The above equation holds, because an approximation for the variance of ratio proposed in Stuart and Ord (1998),

$$\text{Var} \left( \frac{Y}{X} \right) = \frac{E^2 Y}{E^2 X} \left[ \text{Var} (Y) - 2 \frac{\text{Cov}(X,Y) \ E \bar{X}}{E^2 \bar{Y} \ E \bar{X}} + \text{Var}(X) \right].$$

Naturally,

$$\hat{\sigma}_{rj}^2 = \left( \frac{u_Y^r}{u_X^r} \right)^2 + \frac{1}{n_j} \left[ \frac{1}{n_j} - 2 \frac{1}{n_j} \rho \sigma_X \sigma_Y + \frac{1}{n_j} \right]$$
APPENDIX C

Proof of the Equations

\[ Corr(logT, PTR_{wk8}) = \rho \]

\[ Cov(logT, PTR_{wk8}) = \rho \sigma_T \sigma_{PTR} \]

\[ Cov(0.4PTR_{wk8} + \varepsilon_{TD}, PTR_{wk8}) = \rho \sigma_T \sigma_{PTR} \]

\[ 0.4 \sigma^2_{PTR} = \rho \sigma_T \sigma_{PTR} \]

\[ \sigma_{PTR} = \rho \frac{\sigma_T}{0.4} \]

\[ \sigma^2_T = 0.4^2 \sigma^2_{PTR} + \sigma^2_{TD} \]

\[ \sigma^2_T = 0.4^2 \rho^2 \frac{\sigma^2_T}{0.4^2} + \sigma^2_{TD} \]

\[ \sigma^2_{TD} = (1 - \rho^2) \sigma^2_T \]
REFERENCES


FDA Guidance for industry: adaptive design clinical trials for drugs and biologics.


