University of Alberta

COMPARATIVE INVESTIGATION ON CLINICAL TRIAL DESIGNS

by

Jing Wang

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of
Master of Science

in

Biostatistics

Department of Mathematical and Statistical Sciences

Spring, 2012
Edmonton, Alberta

Abstract

A number of interesting subjects relevant to optimality of design, cost efficiency evaluation, and the adaptive treatment allocation for response-adaptive repeated measurement designs have been reviewed and discussed. First we introduce some optimal crossover designs, and compare those designs with completely randomized trials and N-of-1 trials in terms of their relative design efficiency and cost saving, followed by a discussion of three statistical models for repeated measurement designs. Then the response-adaptive design in comparison with standard randomized clinical trials has been elaborated. An adaptive treatment allocation scheme for a multiple-objective response-adaptive repeated measurement design is presented in detail; and the simulation study illustrates how the adaptive treatment allocation scheme works efficiently to simultaneously achieve two objectives: increasing estimation precision and treatment benefit.

Acknowledgements

This thesis would not have been possible without the support of many people. Firstly, I would like to express my gratitude to my supervisor, Dr. Keumhee Carriere Chough who was abundantly helpful and gave me her guidance and support throughout the thesis process. I want to show my thanks to Dr. Sunita Vohra who introduced me the N-of-1 trial design and motivated me to do it with methodological research. I am also greatly thankful to Dr. Pengfei Li, for his precious time and valuable advice for this thesis.

Also thanks to Dr. Yuanyuan Liang for her R codes and Dr. Hongjia Chen for his extensive help with Latex.

Furthermore, I would like to thank the department of Mathematical and Statistical Sciences, University of Alberta, for offering me the opportunity to study here and the financial support.

Finally, thanks to my family members, especially my husband, for supporting and encouraging me to pursue this degree.

Table of Contents

List of Tables

Chapter 1

Introduction

Clinical trials are defined in simple terms by Piantadosi (2005)[16] as experiments that allow researchers to test medical treatments on human subjects. According to Byar et.al (1976)[3], "randomized clinical trials remain the most reliable method for evaluating the efficacy of therapies." The following chapter will briefly review some classic clinical trials commonly applied in medical research. The motivation as well as the organization of my thesis will also be presented.

In a randomized clinical trial (RCT), which is more often called the completely randomized trial in the most statistical literatures, a group of patients are randomly assigned to subgroups - for example, two treatment groups (1 and 2) - such that patients in group 1 are treated with intervention A, and those in group 2, with intervention B. Such a design could achieve unbiased estimates of treatment effects because it eliminates (or at least greatly reduces) the allocation bias in the assignment of treatment, from known or unknown prognostic factors. Although it is recognized as the "gold standard" method for providing evidence on efficacy, there are circumstances in which the design

may fail due to lower power to detect the efficacy of treatment, a difficulty in recruiting required number of subjects, or insufficient fiscal resources. To deal with those situations, many other types of designs with some theoretical as well as clinical advantages have been proposed and examined by clinical investigators and researchers.

A repeated measurement design (RMD) is one of the alternatives. In a RMD, multiple measurements are taken in sequence on the same experiment unit, such as a patient in clinical trials, and the response measurements are usually taken at different times in the trial. There are two general types of the RMDs in clinical trials: parallel group designs and crossover designs (CODs).

In parallel group designs, the subgroups of patients defined by a treatment receive a single therapy and are followed up over time. Table 1 shows an example for a simple two-treatment, two-period parallel group design.

	Period 1	Period 2
Seq.group 1	A	A
Seq.group 2	B	B

Table 1.1: 2-treatment 2-period 2-sequence

In CODs, all patients are given all of the treatments under investigation in a trial. The simplest crossover design is the two-treatment trial with two periods, in which patients are divided into two subgroups, and each patient in group 1 receives intervention A first, and B second, while each in group 2 receives treatment B first and A second. The layouts of a crossover trial with two periods and three periods are presented in Table 2 and Table 3, respectively.

N-of-1 trials are single-patient, double-blind, randomized, and multi-crossover trials, as Zucker (2010) [20] noted. In other terms, the combined N-of-1 trial

	Period 1	Period 2
Seq. group 1	A	B
Seq. group 2	B	A

Table 1.2: 2-treatment 2-period 2-sequence

	Period 1	Period 2	Period 3
Seq. group 1	A	B	A
Seq. group 2	B	A	A

Table 1.3: 2-treatment 3-period 2-sequence

design is a special type of two-treatment repeated-crossover designs for small clinical trials. For example, in a series of N-of-1 trials with six-week treatment periods (one week as a treatment period), each patient undergoes 3 pairs of treatment periods. Each pair of treatment periods includes a treatment A and a treatment B. In the first treatment pair, the treatment allocation is blocked randomized; while in the other two pairs, it is simply randomization. Table 4 gives the layout of a six-period N-of-1 trials design.

	Period 1	Period 2	Period 3	Period 4	Period 5	Period 6
Seq. group 1	A	B	A	B	A	B
Seq. group 2	B	A	B	A	A	B

Table 1.4: 2-treatment 6-period 2-sequences

In our study, We would like to illustrate the optimality properties of the designs that are commonly used in clinical trials and apply some evaluation criteria, such as design efficiency in estimation of treatment effects and cost efficiency comparison, in order to compare those designs. Because of the growing interest in N-of-1 trials and response adaptive designs and their attractive characteristics, we also put them in our investigation.

The statistical models used for these clinical trials designs are similar, but

the models for crossover trials data are usually more general and complex; we will thus focus our discussion on such models without loss of generality. To this end, this paper is organized as follows: Chapter 2 provides a review of the literature detailing current methods of analyzing repeated measures data. Chapter 3 investigates the statistical power and cost efficiency of these interesting designs. Chapter 4 introduces a multiple-objective response-adaptive design and an illustration of the use of an adaptive treatment allocation scheme. Chapter 5 presents data analysis to generate optimal designs for three-period and six-period crossover designs. Chapter 6 summarizes the general results of this research and suggests a few ideas for future research.

Chapter 2

Literature Review

This chapter is composed of two parts. The first part discusses the major issues related to the designs that were introduced in Chapter 1. In the second part, we describe a few general statistical methods commonly applied to those designs, such as the linear mixed model and meta-analysis. Many more statistical models for clinic trials, such as Bayesian hierarchical method, are not covered here in any detail since they are beyond the scope of this thesis. So for the purpose of this thesis, we will focus on the models that we overviewed. For simplicity, we restrict our discussion to the two treatments throughout this thesis.

2.1 General Clinical Trials Designs

We explained the basic definitions of the most common classes of clinical trial designs in the previous chapter. It has been generally accepted that the COD, without some limitations of their own, or when feasible, appears to be a more practical approach than the RCT or parallel group design in clinical trials.

Essentially, this is because it can offer improved efficiency compared to other traditional designs when comparing different treatment effects. Because each patient in a COD receives both treatments, with his or her own control, the sample size required to yield equally precise estimation of the treatment difference is only half as many subjects as in a RCT. Moreover, since the within-subject response measures are usually positively correlated, this allows the sample size to be reduced further.

Consider a simple crossover trial AB/BA. Let μ_A and μ_B be the mean response measures on patients taking treatments A and B, respectively. Assume that both treatment effects are estimated with variance σ^2 and that there are no period or carryover effects. The variance for the estimate of treatment effect difference, $\hat{\tau} = \hat{\mu}_A - \hat{\mu}_B$, is then

$$\begin{aligned} \text{var}(\hat{\tau}) &= \frac{2\sigma^2}{m} - 2\text{cov}(\hat{\mu}_A, \hat{\mu}_B) \\ &= \frac{2\sigma^2}{m}(1 - \rho_{AB}) \end{aligned}$$

based on Piantadosi (2005)[16], where m is the number of subjects in each sequence group, and ρ_{AB} is the response correlation on treatments A and B (assumed to be the same for all individuals).

In RCTs or parallel group designs, ρ_{AB} is zero since different subjects form the two-treatment groups A and B. In CODs, ρ_{AB} is generally expected to be positive because the response measures on the two treatments come from the same subject (Piantadosi, 2005)[16]. Therefore, the COD could be more efficient than a RCT or parallel group design with a smaller variance. This means that the COD has more power to detect a treatment difference with a smaller sample size.

However, a number of disadvantages do arise with CODs. Carryover effect is one of the primary concerns. The treatment effect from one period may persist in the following period; this is also called a residual effect. One potential problem with carryover effects is that the design can not yield unbiased estimates of the treatment effect. Although there are strategies to deal with such difficulties, the most helpful approach is to prevent this effect in the designing stage of a trial by providing a sufficiently long washout period between the treatment periods.

N-of-1 trials have become very attractive to clinical researchers in recent decades. Such a design takes advantage of the principal strength of crossover trials, and emphasizes individual-focused assessment. According to Zucker et.al (1997)[21], a RCT or COD can provide information regarding the relative treatment efficacy for a study population, while an N-of-1 trial can provide information regarding the relative treatment effectiveness for an individual patient. In contrast to the standard trial design, the key advantages of the N-of-1 trial design are that clinical investigators can use individual outcomes to make individual patient treatment decisions and to obtain comparative effectiveness estimates for the general population. On the other hand, they share the similar limitations or shortcomings as usual CODs. For example, such designs may often experience a high dropout, or they may be well suited only to certain disease studies, whose conditions are chronic and stable, require long-term medication and the proposed treatments have a quick onset of action and cease to act soon after they are discontinued, but each aspect needs its own suitability consideration.

In brief, a RCT is the traditional and classic design conducted in clinical trial for its simple and reliable method to test two treatment effects. A COD

can compete against a RCT by its sufficiently increased design efficiency with a smaller sample size, but it also brings some concern of carryover effects. The N-of-1 trial design takes the form of more like a COD, however, some condition or randomization rule is more restricted/complicated in such a design. The most important characteristic that makes the N-of-1 trials different from other designs, is that they are collections of individual trials and intend to identify individual treatment effect, rather than focus on the overall mean treatment effect for a population, which is also the most powerful advantage of such a design.

2.2 Linear Mixed Models for Repeated Measures Data

Let $Y_j = (y_{1j}, y_{2j}, ..., y_{pj})'$ denote the response vector for the jth patient; we assume a general linear mixed model in matrix form is

$$Y_j = X_j\beta + Z_j b_j + \epsilon_j \tag{2.1}$$

where X and Z are design matrices for fixed effects and random effects, respectively, β is the fixed effect vector, b_j is the random effect vector, and ϵ_j is the random error term.

Suppose that the random effects are normally distributed with

$$\mathrm{E}\begin{pmatrix} b_j \\ \epsilon_j \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad \mathrm{Var}\begin{pmatrix} b_j \\ \epsilon_j \end{pmatrix} = \begin{pmatrix} \mathbf{V} & 0 \\ 0 & \mathrm{R} \end{pmatrix}$$

then Y_j follows a multivariate normal distribution, with variance,

$$\Sigma = Z\mathbf{V}Z' + \mathrm{R}$$

The traditional linear mixed model for repeated measurement designs is a specific case of the above model, with the assumptions that the fixed effect set $\beta = (\mu, \pi, \tau, \gamma)'$ comprises the mean treatment effect, period effect, treatment effect, and first-order carryover effect due to treatment given in previous period; and the random effect terms include random subject effects ζ_j and random error effects ϵ_j. Assume that ζ_j and ϵ_j are each independent multivariate normal random vectors with mean zero and variance $\sigma_s^2\mathbf{11}'$ and $\sigma_e^2\mathrm{I}$, respectively.

Many investigators have explored various analytic approaches for analyzing the repeated measures data based on the traditional model. For example, Grizzle (1965) [9] used the classical methods to estimate the treatment effects for an AB/BA design, however, in which the carryover effects was omitted from the model. In face of carryover effects, the unbiased estimate of treatment effect contrast is not estimable unless the data from second period are discarded.

Let $\bar{Y}_{i.k}$ be the average response for all subjects in ith period from sequence group k, where i=1,2, k=1,2, and let τ be the treatment effect contrast, i.e. τ_1-τ_2=τ, subject to τ_1+τ_2=0.

If the assumption of no carryover effects or equal carryover effects is valid, the unbiased estimate of τ can be given by

$$\hat{\tau} = \{(\bar{Y}_{1.1} - \bar{Y}_{2.1}) + (\bar{Y}_{2.2} - \bar{Y}_{1.2})\}/2$$

	Seq.group AB		Seq.group BA	
		mean		mean
Period 1	$\mu + \pi_1 + \tau_1$	$\bar{Y}_{1.1}$	$\mu + \pi_1 + \tau_2$	$\bar{Y}_{1.2}$
Period 2	$\mu + \pi_2 + \tau_2 + \gamma_1$	$\bar{Y}_{2.1}$	$\mu + \pi_2 + \tau_1 + \gamma_2$	$\bar{Y}_{2.2}$

Table 2.1: The layout of design AB/BA

Also let γ_1 and γ_2 be the carryover effects of treatment τ_1 and τ_2 on the first period, respectively; and $\gamma_1 \neq \gamma_2$, then estimation of τ is using the first period,

$$\hat{\tau} = \bar{Y}_{1.1} - \bar{Y}_{1.2}$$

Brown (1980)[2] also discussed this model. One way to deal with such situations, he noted, is combining additional treatment sequences to the design AB/BA; that is, using the design AB, BA, AA, and BB, which was proposed by a few earlier investigators, e.g.Cochran and Cox (1957)[8]. Wallenstein (1979) described another approach for valid estimation with carryover effects in the design AB/BA. The baseline measures preceding each treatment period are required.

On the other hand, the disturbance of carryover is mainly restricted to two-period crossover designs. As Laska et al.(1983)[12] noted, when the number of periods is greater than 2, there is little penalty due to the presence of carryover; even the baseline observations add little value to the statistical power.

The model (2.2) has also prompted numerous discussions about the comparative efficiency and cost-savings of various designs. For example, Brown (1980)[2] presented a simple procedure to evaluate the cost efficiency of a two-treatment COD with sequences AB and BA relative to RCT for the case in which there is no carryover effect. The cost can be separated into two com-

ponents: let S_0 be the cost of recruiting a new subject, and S_1 be the cost of treating and measuring a recruited subject in a given period. Also suppose that these costs S_0 and S_1 are the same for both COD and RCT.

Then the total costs for a COD is S_{co}, with n subjects in each sequence group,

$$S_{co} = 2nS_0 + 4nS_1$$

and the total costs for a RCT is S_{cr}, with m subjects in each treatment group,

$$S_{cr} = 2mS_0 + 2mS_1$$

Thus R, defined as $R{=}S_{co}/S_{cr}$, is the relative cost of these two trial designs when the treatment effects are estimated to be equally precise. Note that the estimates of variance of the treatment effects in these two cases are $\mathrm{Var}(\hat{\tau}_{co}) = \sigma_e^2/n$ and $\mathrm{Var}(\hat{\tau}_{cr}) = 2(\sigma_s^2 + \sigma_e^2)/m$, respectively. The ratio of the variances is therefore expressed as $\mathrm{Var}(\hat{\tau}_{co})/\mathrm{Var}(\hat{\tau}_{cr}) = m\sigma_e^2/2n(\sigma_s^2 + \sigma_e^2)$, and the relationship between required sample size is $n = (m/2)(\sigma_s^2/\sigma_s^2 + \sigma_e^2)$. The relative cost, shown in Brown(1980)[2], can thus be expressed as

$$R_{co/cr} = \frac{1 + 2S_1/S_0}{2(1 + S_1/S_0)}\left(\frac{1}{1 + \sigma_s^2/\sigma_e^2}\right)$$

Without considering carryover effects, the method of cost efficiency evaluation favors CODs. However, if comparing a four-sequence crossover trial to a RCT, it shows that although the a four sequences design is necessary for a two-treatment two-period COD to have an unbiased estimate of treatment difference in the presence of carryover, its economical advantage as a crossover

is defeated (Brown 1980)[2].

Carriere and Huang (2000)[5] extended the discussion on two- and three-period CODs with various model assumptions, and explored the cases for which there is a negative within-subject correlation. Using the same defined relative cost, R, they made a cost efficiency evaluation for those interest designs. They emphasized in their clear discussion that crossover trials performed better than other independent group designs in terms of efficient estimation and cost savings, especially in cases of large recruiting cost (small S_1/S_0) and large σ_s^2/σ_e^2.

2.3 Linear Mixed Models for N-of-1 trials

A much simple form of linear mixed model is usually used for N-of-1 trials data. Recall that N-of-1 trials design can be seen as a special individual-focused clinical crossover trials.

Considering a situation where the outcome depends only on treatment, the model (2.1) can be written as,

$$y_j = \mu + \zeta_j + \epsilon_j \tag{2.2}$$

where μ is the average treatment effect, $\mu + \zeta_j$ specify the individual treatment effect due to subject j, and ϵ_j specifies other sources of errors between outcomes. The model (2.3) may be regarded as a special case of the model (2.1) with $X_j = Z_j = 1$, $\beta = \mu$ and $b_j = \zeta_j$. If ζ_j is assumed to be a constant

across subjects, Equation (2.3) can be further simplified as

$$y_j = \mu + \epsilon_j \tag{2.3}$$

The only source of variation in response comes from the within-subject variance, ε_w^2. Such a model is called a fixed effect model.

Rather than assuming that the treatment effect is fixed, we allow that the individual treatment effect depends on many patient-specific characteristics , i.e, ζ_j has a normal distribution with mean 0 and variance ε_b^2. A random effect model can thus be obtained by

$$\begin{aligned} y_j &= (\mu + \zeta_j) + \epsilon_j \\ &= \mu_j + \epsilon_j \end{aligned} \tag{2.4}$$

The method of the random effect analysis is to decompose the variance into two component parts: between-subject variance ε_b^2 and within-subject variance ε_w^2.

Another common method used for N-of-1 trials is the summary meta-analysis. The approach first fits a model to the individual-patient data from each subject separately, then combines the individual treatment effect estimates, y_j, using a meta-analytical approach. There is no essential difference between meta-analysis and the linear mixed model in terms of methodologies. However, meta-analysis provides another way to pool individual treatment effects in estimating the average treatment effects. Note that in meta-analysis, each subject's trial is treated as an independent study.

If we assume that all the studies share the true treatment effect, μ, then

the observed effects y_j for individual study will be distributed about μ, only with a within-study variance ε_w^2; and the resulting model will have a same form as Equation (2.4) for the fixed effect model analysis.

However, there is generally no reason to assume that each study has exactly the same true mean effect in meta-analysis. Therefore, a random effect model that makes allowance for an additional between-studies variance is preferable. By the same token, in N-of-1 trials each single subject can have a different effect, μ_j; thus we are back to Equation (2.5) for the random effect model.

In meta-analysis, each observed treatment effect, y_i is weighted by the inverse of its estimated variance, then a weighted average treatment effect is obtained by

$$\hat{\mu} = \sum_{j=1}^{n} \omega_j y_j / \sum_{j=1}^{n} \omega_j \tag{2.5}$$

The difference between the fixed and random effect model analysis for meta-analysis is that the variance in random effect model includes one more component, ε_b^2, the between-subject variance. Therefore the weights, w_j, used in the formulation, would have a corresponding change in the two models.

Linear mixed models and meta-analysis should provide identical results since they share the same model (2.1),$Y_j = X_j\beta + Z_j b_j + \varepsilon_j$. The generalized least squares estimate for β is given by $\hat{\beta} = \left(\sum_1^n X_j'\Sigma^{-1}X_j\right)^{-1}\left(\sum_1^n X_j'\Sigma^{-1}y_j\right)$, where Σ is a known variance-covariance of Y_j. Since the meta-analytic method is a specific case of linear mixed model with assumption of $X_j = Z_j = 1$, the least squares estimate becomes $\hat{\beta} = \left(\sum_1^n \sigma_j^{-2}\right)^{-1}\left(\sum_1^n \sigma_j^{-2}y_j\right)$. When substituting σ_j^2 with $1/w_j$, we have the same formula for the mean or average treatment effect as Equation (2.6).

Chapter 3

Optimal Design and Cost Efficiency Comparison

In this chapter, we introduce some general optimal crossover trials, describe how to calculate an information matrix for such designs using Carriere and Reinsel's method (1992)[7], and present the analytical formulae for the interest quantities used in the analysis of precise estimation and cost efficiency comparison. The optimal design here means, a design by which we can acquire a best linear unbiased estimator (BLUE) for treatment effect difference with minimal variance over all possible assignments of N patients to the 2^p possible sequence groups. In general, optimal designs differ for the various models.

3.1 Common Optimal Crossover Designs

For p=3, there are a total of eight possible sequence groups available in crossover designs: AAA, ABA, AAB, ABB and their duals. The dual of a treatment sequence is defined as a new treatment sequence whose treatment

order is inverse of the original sequence's, for example, BAB is the dual of ABA. The design ABB /BAA is optimal over a class of tree-period designs under the traditional model (Laska and Meisner, 1985[13]). In addition, Carriere (1994)[4] describe a nearly optimal three-period design consisting the sequences ABB, AAB and their duals. In such a design, the sequence groups ABB and BAA have a higher allocation proportion of patients than treatment sequences AAB and BBA so that it remains competitive efficiency with various model assumptions.

As Laska and Meisner (1985)[13] discussed, for p=4, the design with the two sequence groups ABBA, AABB and their duals is optimal. For p=6, there are three designs available can achieve the most efficiency of estimating the treatment effects, (1) AABBBA, ABBAAB and their duals; (2) ABBAAB, ABBBAA and their duals; (3) the combination of sequences groups (1) and (2).

As we see when the period is extending, the possible sequences rapidly increase (total 2^p available). However, it is not necessary to apply all of them in a practical setting. For instance, the most common treatment sequences used in N-of-1 trials include only eight groups if we strictly follow its initial treatment plan. Since the treatment allocation rule requires each pair of treatment periods to consist of two different interventions, some sequences like AAABBA or AAAABA may be less likely given in N-of-1 trials. Also in clinical setting, such sequence groups repeating the same treatment on patients may be ethically debatable if the treatment efficacy is unknown. In such cases, we need to consider the issues arising from balancing the statistical optimality and clinical suitability.

We will examine the following designs in detail in terms of different effi-

ciency criteria.

Design I	ABB/BAA
Design II	ABB,AAB /BAA,BBA
Design III	ABBA,AABB and their duals
Design IV	ABBAAB,AABBBA and their duals
Design V	ABABAB,ABBAAB,ABABBA,ABBABA and their duals

Table 3.1: Two-treatment multiple-period competing designs

3.2 Information Matrix

Recall the traditional linear mixed model (2.1), we can rewrite it using the following form:

$$y_{ijk} = \mu + \pi_i + \tau_{d(i,k)} + \gamma_{d(i-1,k)} + \zeta_{jk} + \epsilon_{ijk}$$

with $\zeta_{jk} \sim N(0, \sigma_s^2)$ and $\epsilon_{ijk} \sim N(0, \sigma_e^2)$. Here, y_{ijk} denotes the responses of subject j from treatment sequence group k in period i, μ is the mean effect, π_i is the ith period effect, where i=1,2, ...,p. $d(i,k)$ denotes the treatment given in period i of sequence k, k=1,2, ...,2^p, then $\tau_{d(i,k)}$ is the direct effect of treatment used in period i from sequence k, and $\gamma_{d(i-1,k)}$ is the carryover effect in period i-1 of sequence k.

Let $\tau = (\tau_A - \tau_B)/2$ and $\gamma = (\gamma_A - \gamma_B)/2$, with constraints $\tau_A + \tau_B = \gamma_A + \gamma_B = 0$. Equivalently, $\tau = \tau_A = -\tau_B$ and $\gamma = \gamma_A = -\gamma_B$. Then we have the following model:

$$y_{ijk} = \mu + \pi_i + \mathrm{I}_{d(i,k)}\tau + \mathrm{I}_{d(i-1,k)}\gamma + \zeta_{jk} + \epsilon_{ijk} \tag{3.1}$$

where

$$\mathrm{I}_{d(i,k)} = \begin{cases} 1 & \text{if } d(i,k) = \text{A} \\ -1 & \text{if } d(i,k) = \text{B} \end{cases},$$

and

$$\mathrm{I}_{d(i-1,k)} = \begin{cases} 0 & \text{if } i = 1 \\ 1 & \text{if } i > 1 \text{ and } d(i-1,k) = \text{A} \\ -1 & \text{if } i > 1 \text{ and } d(i-1,k) = \text{B} \end{cases}$$

Both the random subject effect ζ_{jk} and the random error ϵ_{ijk} are assumed to be independently, identically distributed with mean zero and variance σ_s^2 and σ_e^2, respectively. Further, ζ_{jk} and ϵ_{ijk} are mutually independent, for all responses from different subjects. Then the covariance-variance structure of the p-vectors y_{jk} is considered as,

$$\mathrm{cov}(y_{ijk}, y_{i'j'k'}) = \begin{cases} \sigma_s^2 + \sigma_e^2 & \text{if } i \neq i', j = j' \text{ and } k = k' \\ \sigma_s^2 & \text{if } i \neq i', j = j' \text{ and } k \neq k' \\ 0 & \text{otherwise} \end{cases}$$

or in matrix form,

$$\mathrm{cov}(y_{jk}) = \Sigma = \sigma_e^2 \mathrm{I}_p + \sigma_s^2 \mathbf{1}_p \mathbf{1}_p' \tag{3.2}$$

and the correlation coefficient between y_{ijk} and $y_{i'jk}$ is

$$\rho = \begin{cases} \frac{\sigma_s^2}{\sigma_s^2 + \sigma_e^2} & \text{if } i \neq i' \\ 1 & \text{if } i = i' \end{cases}$$

Define $N = \sum_k n_k$, and $y_{.k} = \sum_{j=1}^{n_k} y_{jk}$, where j=1,2,...,n_k, then,

$$\Sigma^{-1} = \sigma_e^{-2} \left\{ \mathrm{I}_p - \frac{\rho}{1+(p-1)\rho} \mathbf{1}_p \mathbf{1}_p' \right\}$$

We can also have the model (3.1) in matrix form:

$$y_{jk} = X_k \beta + \mathbf{1}_p \zeta_j + \epsilon_{jk} \tag{3.3}$$

where $\beta = (\mu, \pi_2, ..., \pi_p, \tau, \gamma)'$ is a vector of unknown parameters and X_k is a $p \times (p+2)$ design matrix for sequence k. For instance, if p=6 and the treatment sequence group k, is ABABAB, then the corresponding design matrix is,

$$X_k = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & -1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & -1 \\ 1 & 0 & 0 & 1 & 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & -1 \\ 1 & 0 & 0 & 0 & 0 & 1 & -1 & 1 \end{pmatrix}$$

For the dual balanced designs, by dual lemma (Laska and Meisner, 1985)[13], the sufficient conditions for optimal designs are $X_k^* = -X_{k'}^*$ and $n_k = n_{k'}$, where k and k' are dual sequences, and X_k^* refer to the columns that may be negatives on the corresponding columns of their dual sequences, such as X_k^τ and X_k^γ. For the remaining columns in X_k, like $X_k^\pi = (X_k^{\pi_2}, X_k^{\pi_3}, ..., X_k^{\pi_p})$, they are same for all sequence groups. Therefore we can also have $\sum_1^k n_k X_k^\pi \Sigma^{-1} X_k^* = 0$ for all k.

Using the above notations, Carriere and Reinsel (1992)[7] gave a best un-

biased estimator of treatment effect differences, $\beta_{\tau,\gamma} = (\tau, \gamma)'$, based on generalized least squares methods,

$$\hat{\beta}_{\tau,\gamma} = \left(\sum_{1}^{k} n_k X_k^{*'} \Sigma^{-1} X_k^* \right)^{-1} \left(\sum_{1}^{k} X_k^{*'} \Sigma^{-1} y_{.k} \right)$$

and

$$\mathrm{cov}(\hat{\beta}_{\tau,\gamma}) = \left(\sum_{1}^{k} n_k X_k^{*'} \Sigma^{-1} X_k^* \right)^{-1}$$

Note that the information matrix for $\hat{\beta}_{\tau,\gamma}$ is,

$$\mathrm{I}(\tau, \gamma) = \left(\sum_{1}^{k} n_k X_k^{*'} \Sigma^{-1} X_k^* \right)$$

Recall that $d(i,k)$ denotes the treatment used in the ith period of the kth sequence. Denote a_{ik}=1 when $d(i,k)$=A and -1 when $d(i,k)$=B, the information matrix for each sequence k can also be given as,

$$\begin{aligned} \mathrm{I}(\tau, \gamma) &= \frac{\sum n_k}{\sigma_e^2} \begin{pmatrix} p - ct_k^2 & g_k - ct_k \tilde{t_k} \\ g_k - ct_k \tilde{t_k} & (p-1) - c\tilde{t_k^2} \end{pmatrix} \\ &= \frac{\sum n_k}{\sigma_e^2} \begin{pmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{pmatrix} \end{aligned} \tag{3.4}$$

as shown in Carriere and Reinsel (1992)[7], where c=$\rho/[1+(p-1)\rho]$, g_k=$\sum_{i=1}^{p-1} a_{ik} a_{i+1,k}$, t_k=$\sum_{i=1}^{p} a_{ik}$, and $\tilde{t_k}$=$\sum_{i=1}^{p-1} a_{ik}$. So the optimal design giving the minimum of variance of treatment effect, $var(\tau)$, is the one that maximizes the information matrix value given in Equation (3.4). With this procedure, we can assess any dual balanced designs in regard to their efficiency in estimating treatment effect, τ.

According to the above information matrix, we have

$$\text{var}(\hat{\tau}) = \frac{\sigma_e^2}{\sum n_k}\frac{I_{22}}{D}$$

where $I_{22}= (p-1) - c\tilde{t_k^2}$ and $D= I_{11}I_{22} - I_{12}^2$. Note that $\text{var}(\hat{\tau})$ is dependent on the value of ρ.

3.3 Efficiency of Estimating Treatment Contrast

Suppose in a dual balanced design, we have $n_k=\frac{1}{s}N$ subjects in the treatment sequence k, where k=1,2, ...,s sequences, and $N=\sum n_k$. Also suppose the repeated measures model has the random subjects effects under a completely symmetric covariance structure, $\text{cov}(y_{jk})=\sigma_e^2\text{I}_p + \sigma_s^2\mathbf{1}_p\mathbf{1}_p'$.

Design I

Design I is the optimal trial within the class of three-period RMDs. Given $c=\rho/(1+2\rho)$, $g=(0,0)'$, $t=(-1,1)'$, $\tilde{t}=(0,0)'$,

$$\begin{aligned}\text{var}(\hat{\tau}) &= \frac{\sigma_e^2}{n_1 * \frac{D}{I_{22}} + n_2 * \frac{D}{I_{22}}} \\ &= \frac{\sigma_e^2}{N} \times \frac{1+2\rho}{3+5\rho}\end{aligned}$$

Design II

To give illustration but simplify calculation, we consider the case with equal number of subjects included in each sequences for design II, that is, $n_k=\frac{1}{4}N$.

Note that the nearly optimal design thus actually has smaller $var(\hat{\tau})$ than the following estimate,

$$\begin{aligned} \text{var}(\hat{\tau}) &= \frac{\sigma_e^2}{n_1 * \frac{D}{I_{22}} + n_2 * \frac{D}{I_{22}}} \\ &= \frac{\sigma_e^2}{N} \times \frac{1+2\rho}{3+5\rho-\rho^2} \end{aligned}$$

Design III

Design III is optimal for p=4. We have c=$\rho/(1+3\rho)$, g=$(-1,-1,1,1)'$, t=$(0,0,0,0)'$, $\tilde{t}$=$(-1,1,1,-1)'$, thus $var(\hat{\tau})$ for the design ABBA, AABB / BAAB, BBAA is given by,

$$\begin{aligned} \text{var}(\hat{\tau}) &= \frac{\sigma_e^2}{\sum_{k=1}^{4} n_k * \frac{D}{I_{22}}} \\ &= \frac{\sigma_e^2}{4N} \end{aligned}$$

Note that the variance estimation of τ do not depend on ρ in this case.

Design IV

For this six-period optimal RMD, given c=$\rho/(1+5\rho)$, g=$(-1,1,-1,1)'$, t=$(0,0,0,0)'$, and $\tilde{t}$=$(1,-1,-1,1)'$, we have

$$\begin{aligned} \text{var}(\hat{\tau}) &= \frac{\sigma_e^2}{\sum_{k=1}^{4} n_k * \frac{D}{I_{22}}} \\ &= \frac{\sigma_e^2}{6N} \end{aligned}$$

Again, the estimate of var(τ) is not dependent upon ρ.

Design V

There are usually eight different sequence groups available in N-of-1 trials. Due to calculation simplicities, we still consider the cases of dual balanced designs. Therefore, we can obtain that g=$(-5,-1,-3,-3,-5,-1,-3,-3)'$, t=$(0,0,\ldots,0)'$, and $\tilde{t}$=$(1,1,-1,-1,-1,-1,1,1)'$ with c=$\rho/(1+5\rho)$; further, the estimate of var(τ) is

$$\begin{aligned} \text{var}(\hat{\tau}) &= \frac{\sigma_e^2}{\sum_{k=1}^{8} n_k * \frac{D}{I_{22}}} \\ &= \frac{\sigma_e^2}{N\left[6 - \frac{11(1+5\rho)}{5+24\rho}\right]} \end{aligned}$$

In addition, we have the estimate of var(τ) for a RCT, or also called completely randomized design, shown in previous chapter, var($\hat{\tau}_{cr}$) = $2(\sigma_s^2+\sigma_e^2)/m$, where $2m$=N. The relative efficiency of two designs is defined as the ratio of their unbiased estimators' variance. Therefor we can evaluate the relative efficiency among designs with their estimates of variances. For example, under the assumption of estimating the treatment effect equally precise, the relative efficiency between design I and completely randomized design is defined as var($\hat{\tau}_{designI}$)/var($\hat{\tau}_{cr}$).

Table 3.2 summarizes the comparison of three optimal CODs, i.e. Design I, III, IV, to RCT, in terms of design efficiency. All of the optimal designs are under the traditional linear mixed model with the equicorrelated covariance structure. Besides the optimal CODs, two more designs of interest are included: 1) Design II, i.e. the nearly optimal design ABB, AAB and their duals; 2) Design V, i.e. the N-of-1 trial design with eight different sequence groups.

The entries are the ratios for such design comparisons, which represent the relative efficiency between designs, with smaller values meaning higher efficiency. As it shows in Table 3.2, for each COD compared with RCT, the ratio values get smaller as ρ increases; and for each level of ρ, the values also become smaller as the number of periods is bigger. It concludes that 1) a design with the period extending becomes more efficient. For example, the four-period optimal COD is more efficient than the three-period optimal COD; and 2) the relative efficiency is enhanced with ρ being larger. Therefore, a design with ρ= 0.8 is more efficient than itself when ρ=0.2. Note that the results from Design I and II are almost the same for their very close formula of the estimate of var(τ). For Design III and V, the similar results suggest such two designs have comparative design efficiency, or in other words, the N-of-1 design with additional period or treatment sequence information, could not improve efficiency compared to the four-period optimal COD.

	Design	Period	$\text{var}(\hat{\tau}_{CO})/\text{var}(\hat{\tau}_{CR})$		
			ρ=0.2	ρ=0.5	ρ= 0.8
I	ABB/BAA	3	0.070	0.045	0.019
II	ABB, AAB/BAA, BBA	3	0.071	0.048	0.020
III	ABBA, AABB/BAAB, BBAA	4	0.050	0.031	0.013
IV	ABBAAB, AABBBA BAABBA, BBAAAB	6	0.033	0.021	0.008
V	ABABAB, ABBAAB, ABABBA ABBABA and their duals	6	0.053	0.033	0.013

Table 3.2: Relative efficiency of competing designs

3.4 Cost Efficiency Comparison

Table 3.3 presents the results of relative cost (R) between each repeated measurement design and RCT. Note that the comparison is for the average patients in the population and no individual patient efficiency evaluation is possible. We assume that all of the conditions related to $R{=}S_{co}/S_{cr}$, discussed in chapter 2, are satisfied here. For each pair of compared designs, let the ratio of S_1/S_0=1/10,1/4,1,4,10, respectively, as well as the ratio of σ_s^2/σ_e^2=1/4,1,and 4, respectively. Similarly as before in Table 3.2, the smaller values of entries (R) in Table 3.3, the more cost efficient or larger cost saving between designs. Especially, a value of 1, means the compared design has same cost as a RCT.

In Table 3.3, none of the values is even close to one. For each COD in comparison to RCT, the value of relative cost becomes smaller as the ratio of σ_s^2/σ_e^2 is large or the ratio of S_1/S_0 is small. The findings suggest that the cost efficiency favors crossover trials, especially when the designs have large recruiting cost (small S_1/S_0) and large σ_s^2/σ_e^2. In general, the cost saving increases as the period of an optimal design extends. This is because the optimal design efficiency can be improved with more periods, under the assumption of other conditions, such as the total number of subjects, N, the values of S_1, S_0, and ratio of S_1/S_0 keep the same. Note that design I and II have almost the same values if we round the numbers at the second decimal points; however, Design V have a comparable cost saving only when the the ratio of S_1/S_0 is less than 1, compared to Design II.

In summary, the multi-period crossover designs achieve great efficiency and cost savings. With additional information on periods, the optimal designs increase both efficiency and cost savings. As Carriere and Huang (2000)[5]

	Design I (t2p3s2)			Design II ($t2p3s4$)			Design III ($t2p4s4$)			Design IV ($t2p6s4$)			Design V (t2p6s8)		
S_1/S_0	$\sigma_s^2/\sigma_e^2 = \frac{1}{4}$,	1,	4												
1/10	0.083	0.054	0.022	0.084	0.056	0.024	0.070	0.044	0.017	0.050	0.031	0.013	0.077	0.049	0.020
1/4	0.098	0.064	0.026	0.099	0.067	0.029	0.088	0.055	0.022	0.069	0.043	0.017	0.107	0.067	0.027
1	0.140	0.091	0.037	0.141	0.095	0.041	0.137	0.086	0.034	0.121	0.076	0.030	0.186	0.117	0.047
4	0.182	0.118	0.048	0.184	0.124	0.053	0.186	0.117	0.047	0.173	0.108	0.043	0.266	0.167	0.067
10	0.197	0.128	0.052	0.199	0.134	0.058	0.204	0.128	0.051	0.191	0.120	0.048	0.295	0.186	0.074

Table 3.3: Relative cost efficiency of competing designs

noted, large recruiting cost (small S_1/S_0) and large σ_s^2/σ_e^2 are important factors to determine the cost effectiveness of crossover trials.

Chapter 4

Response-Adaptive Repeated Measurement Design

This chapter briefly introduces the key characteristics of the response-adaptive designs. Then a randomized treatment allocation scheme to construct a multiple-objective response-adaptive repeated measurement design (RARMD) is reviewed in the second section, followed by application of such an allocation scheme for three, four, and six-period two-treatment repeated measurement designs.

4.1 Background

In general, the designs we discussed in previous chapters, such as RCTs, crossover designs or N-of-1 trials use some similar randomization strategies. For example, in a traditional RCT testing two treatments, patients are equally assigned to the two treatment groups with simple (50-50) randomization. Such an equal allocation scheme is made and fixed in advance of the initiation of

the trial and randomization tends to balance the known or unknown factors among treatment groups and thus guarantees the validity of a statistical inference from the study. For this reason, RCTs are generally considered the most reliable methods to compare treatment effects. However, such clinical trials may cause concern about ethical problems when being applied in practice. That is because half patients have to be exposed to the inferior or unbeneficial treatment group.

In contrast, a response-adaptive design may overcome such a dilemma with a flexible allocation scheme. A response-adaptive design can be defined as a design that makes use of the accumulating information for assigning the best treatment to the most patients in a clinical trials. During the trial, the treatment allocation scheme is adjusted in order to minimize the number of patients assigned to inferior treatment in the trial while providing meaningful statistical inferences.

The adaptive allocation procedures modify the probability of assigning a new patient to a particular treatment based on previously observed responses of patients and allow more patients to be allocated to the potentially beneficial treatment. In addition, these modifications may potentially help reduce the sample sizes required in the trial. Thus the RARMD can offer significant ethical advantages and cost saving over traditional RCTs because of the advantages of unequal treatment allocation.

In a typical RARMD, one adjusts the treatment allocation to fulfill a single objective such as maximizing the number of patients assigned to the better treatment group (Zelen,1969; Wei and Durham,1978), reducing the sample size in a trial (Armitage,1975), or increasing estimation precision of a treatment effect (Kushner,2003). Recently, Liang and Carriere (2009)[15] developed a

multiple-objective RARMD, not only potentially preventing exposing patients to inferior treatment, but also enhancing precision of estimates of parameters. In the following section, we will review the adaptive treatment allocation scheme to construct a multiple-objective RARMD.

4.2 Review of Adaptive Treatment Allocation Scheme

Consider a multiple-objective RARMD with two treatments, in which we desire a randomized allocation scheme to simultaneously serve a dual objective as precision estimation and minimizing patients to unfavorable treatment. We can use a function of the information matrix and an evaluation function to measure the two goals, respectively.

Liang and Carriere (2009)[15] defined a selection criterion as follows,

$$\Phi = \lambda \frac{\Delta(\hat{\mathrm{I}}^k_{j+1})}{\Delta(\hat{\mathrm{I}}^{k^A}_{j+1})} + (1-\lambda)\frac{f_{j,k}}{f_{j,k^B}} \tag{4.1}$$

where $\hat{\mathrm{I}}^k_{j+1}$ represents the estimated information matrix, given the information of the first j patients' responses, and the assumption that the $(j+1)$th patient will be given treatment sequence k. $\Delta(.)$ can be any optimality criterion, for instance, the determinant, the trace or the maximum eigenvalue of the information matrix, $\hat{\mathrm{I}}^k_{j+1}$. And $f_{j,k}$ is an evaluation function for treatment sequence k based on the responses of the first j patients. k^A denotes the treatment sequence that maximizes $\Delta(.)$, and k^B denotes the the treatment sequence that maximizes $f_{j,k}$. k^A and k^B are not necessarily the same.

In Equation (4.1), the first part $\frac{\Delta(\hat{\mathrm{I}}^k_{j+1})}{\Delta(\hat{\mathrm{I}}^{k^A}_{j+1})}$ works to find a treatment sequence which could maximize the information matrix, while the second part $\frac{f_{j,k}}{f_{j,k^B}}$ intends to detect a treatment sequence which perform best given the first j patients' data. Thus a lagrange coefficient λ, a constant between zero and one, can weight and balance the two objectives. For example, if λ takes one, the resulting treatment sequence will achieve most precise estimation but sacrifice the ethical advantages. If λ is set to zero, we may only concern the efficacy of the treatment sequences (Liang and Carriere, 2009)[15]. Therefore, By giving a value of λ prior to the trial, we can choose a treatment sequence, satisfying the two objectives to a degree and maximizing the value of Φ, as the best treatment sequence for the incoming $(j+1)$th patient.

4.3 Application of Adaptive Allocation Scheme

This part examines the adaptive treatment allocation scheme when it is applied to the two-treatment repeated measurement designs with dichotomous or continuous outcomes.

Recall the model (3.3) for a repeated measurement design,

$$y_{jk} = X_k\beta + \mathbf{1}_p\zeta_j + \epsilon_{jk}$$

A self and mixed carryover effects model based on the above model for a multiple-objective RARMD can be written as

$$y_{ijk} = \mu + \pi_i + \tau_{d_k(i,j)} + \mathrm{I}_{ijk}\gamma_{d_k(i-1,j)} + (1-\mathrm{I}_{ijk})\delta_{d_k(i-1,j)} + \zeta_{jk} + \epsilon_{ijk} \qquad (4.2)$$

where y_{ijk} is the response variable for observation on subject j in ith period from sequence group k, $d_k(i,j)$ is the treatment given to subject j in period i for sequence k, and I_{ijk} is a 0,1 dummy variable where it gives a value of 1 if $d_k(i,j) = d_k(i-1,j)$ or a 0, otherwise.

In the fixed effects set $\beta=(\mu,\pi,\tau,\gamma,\delta)'$, μ is the overall mean, π is the period effect, and τ is the direct treatment effect, for its coefficient is 1 if given treatment A, or -1 if B. γ is defined as the self carryover effect with coefficient, taking 1 if treatment A precedes treatment A and -1 if treatment B precedes treatment B; and δ is the mixed carryover effect with coefficient, equaling 1 if treatment A precedes treatment B and -1 if treatment B precedes treatment A. Both γ and δ have coefficients as zero in the first period.

The random subject effect ζ_{jk} and random error ϵ_{ijk} are still assumed to be independently and multi-normally distributed, with $N \sim (0,\sigma_s^2)$ and $N \sim (0,\sigma_e^2)$, respectively.

4.3.1 Adaptive Allocation Scheme for Repeated Measurement Design with Dichotomous Responses

If the response in a multiple-objective RARMD is dichotomous, for example, success or failure, patients' responses can be assumed to be independent and identically distributed with the Bernoulli distribution. Suppose $\delta_{ti} = (\delta_{ti1},...,\delta_{tiJ})$ is the treatment t given in period i, where t= A or B, such that δ_{tij}=1 if the jth patient receiving treatment t in period i and δ_{tij}=0 otherwise, and y_{tij} is the corresponding response. Then $m_{ti} = \Sigma_{j=1}^{J}\delta_{tij}$ denotes the number of patients receiving treatment t in period i, while $S_{ti} = \Sigma_{j=1}^{J}y_{tij}\delta_{tij}$ represents the number of success in period i for treatment t.

After j patients have been treated in the trial, the likelihood function is

$$\begin{aligned} L_j &= \Pi_t \Pi_{i=1}^{p} \Pi_{j=1}^{J} [\nu_{ti}^{y_{tij}} (1-\nu_{ti})^{1-y_{tij}}]^{\delta_{tij}} \\ &= \Pi_t \Pi_{i=1}^{p} [\nu_{ti}^{S_{ti}} (1-\nu_{ti})^{m_{ti}-S_{ti}}] \end{aligned} \tag{4.3}$$

where $\nu_{ti} = (\nu_{A1}, \nu_{A2}, ..., \nu_{Ap}, \nu_{B1}, ..., \nu_{Bp})'$ is the probability of success on period i for treatment A or B.

For example, in a trial with two periods, Equation (4.3) becomes

$$\begin{aligned} L_j \propto \nu_{A1}^{S_{A1}} (1-\nu_{A1})^{m_{A1}-S_{A1}} \times \nu_{A2}^{S_{A2}} (1-\nu_{A2})^{m_{A2}-S_{A2}} \\ \times \nu_{B1}^{S_{B1}} (1-\nu_{B1})^{m_{B1}-S_{B1}} \times \nu_{B2}^{S_{B2}} (1-\nu_{B2})^{m_{B2}-S_{B2}} \end{aligned}$$

where ν_{A1} is the success probability of treatment A in the first period, S_{A1} is the number of patients receiving treatment A with good results, m_{A1} is the total number of patients receiving treatment A, therefore the first term in above equation represents the likelihood of success for treatment A in the first period, the second term then represents the likelihood of success for treatment A in the second period, and so on.

The log-likelihood function follows

$$log(L_j) = \sum_{t} \sum_{i=1}^{p} [S_{ti} log \nu_{ti} + (m_{ti} - S_{ti}) log(1-\nu_{ti})]$$

Equivalently, if we set l=1,2,...,2p, then S_l=S_{A1}, ..., S_{Ap}, if l=1,..,p, and S_l=S_{B1}, ..., S_{Bp}, if l=$p+1$,...,2p. The above equation can be written as below,

$$log(L_j) = \sum_{l=1}^{2p} [S_l log \nu_l + (m_l - S_l) log(1-\nu_l)]$$

where S_l denotes the lth element of S, $S = (S_{A1}, ..., S_{Ap}, S_{B,P+1}, ..., S_{B,2p})'$, the set of success number with treatment A and B in each period. For instance, $S_{B,p+1}$ means the number of success for treatment B in period 1. ν_l and m_l are the corresponding success probability and number of patients, respectively, i.e. $\nu = (\nu_{A1}, ..., \nu_{Ap}, \nu_{B,P+1}, ..., \nu_{B,2p})'$ and $m = (m_{A1}, ..., m_{Ap}, m_{B,P+1}, ..., m_{B,2p})'$. Then the expected information matrix up to the jth patient, is a $l \times l$ diagonal matrix,

$$I_j = Diag[\mathrm{E}(\frac{S_l}{\nu_l^2} + \frac{m_l - S_l}{(1-\nu_l)^2})]$$

Under the assumption that $(j+1)$th patient will be given treatment sequence k, the estimated information matrix is then

$$I_{j+1} = Diag[\mathrm{E}(\frac{S_l'}{\nu_l^2} + \frac{m_l' - S_l'}{(1-\nu_l)^2})]$$

as shown in Carriere and Liang (2010)[6], where S_l'=S_l+$Diag(\nu \times d_k)$, ν=$(\nu_1, ...\nu_{2p})$, and d_k is an indicator variable set of ν for the kth treatment sequence. As an example, d=(1,0,1,0,1,0) corresponds to the sequence ABA in a three-period RARMD. For a dichotomous response case, an evaluation function $f_{j,k}$ can be defined as the average number of success for treatment sequence k,

$$f_{j,k} = \frac{d_k \times S}{n_k}$$

where n_k=$d_k \times m$, the number of patients in sequence k, and m=$(m_1, ..., m_{2p})'$, the vector of the number of patients in each period for treatment A and B.

4.3.2 Adaptive Allocation Scheme for Repeated Measurement Design with Continuous Responses

If a multiple-objective RARMD based on Model (4.2), has continuous responses, recall that the covariance-variance structure of the y_{jk}, Σ, is given in Equation (3.2),

$$\text{cov}(y_{jk}) = \Sigma = \sigma_e^2 \mathrm{I}_p + \sigma_s^2 \mathbf{1}_p \mathbf{1}_p'$$

For example, in a six-period case, then $\text{cov}(y_{jk}) = \sigma_e^2 \mathrm{I}_6 + \sigma_s^2 \mathbf{1}_6 \mathbf{1}_6'$. Suppose there are four different sequence groups available in the trial, that is, k={ABAABA, ABBBAA, BABBAB, BAAABB }. The design matrix X_k for $\beta=(\mu, \pi, \tau, \gamma, \delta)'$, given the treatment sequence k, can be defined as follows,

$$X_{ABAABA} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & -1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & -1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & -1 \end{pmatrix}$$

$$X_{ABBBAA} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & -1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & -1 & -1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & -1 & -1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & -1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}$$

$$X_{BABBAB} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 1 & 0 & 1 & 0 & 0 & 0 & -1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & -1 & -1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & -1 \\ 1 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 1 \end{pmatrix}$$

$$X_{BAAABB} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & -1 & -1 & 0 \end{pmatrix}$$

Therefore, the estimated information matrix based on the responses of the first j patients, I_j, is

$$\hat{\mathrm{I}}_j \;=\; \sum n_k X_k' \hat{\Sigma}^{-1} X_k \tag{4.4}$$

or in our six-period example, it becomes $(\sum n_k X_k'[\hat{\sigma}_e^2 \mathrm{I}_6 + \hat{\sigma}_s^2 \mathbf{1}_6 \mathbf{1}_6']^{-1} X_k)$, where n_k is the number of patients assigned to treatment sequence k, k=1,2,...,s.

With the assumption that the $(j+1)$th patient taking the treatment sequence k, the above estimated information matrix will become,

$$\hat{\mathrm{I}}_{j+1}^k = \hat{\mathrm{I}}_j + X_k' \hat{\Sigma}^{-1} X_k \tag{4.5}$$

For continuous outcomes, we can assume the sequence group offering the

best treatment will be the one with the largest values of an evaluation function $f_{j,k}$. To choose the treatment sequence for the $(j+1)$th patient, we maximize the selection criterion Φ, Equation (4.1),

$$\Phi = \lambda \frac{\Delta(\hat{\mathrm{I}}^{k}_{j+1})}{\Delta(\hat{\mathrm{I}}^{k^A}_{j+1})} + (1-\lambda)\frac{f_{j,k}}{f_{j,k^B}}$$

Eventually, the total n patients could be assigned by repeating the same procedure.

4.4 Simulation Results

We use a simulation study to investigate the adaptive treatment allocation scheme in various RARMDs with dichotomous and continuous responses, respectively. The number of simulations to be performed in each case is L=1000 with different seeds. Some parameter setting refers to Carriere and Liang (2010)[6], where it is appropriate. In each simulation, suppose that λ=1,0.9,0.7,0.3, and 0, and the total number of patients in a trial, N=12(10), 40, 80, and 100, respectively.

4.4.1 Three-Period Repeated Measurement Design

In a two-treatment three-period RMD, eight different treatment sequences are available. We consider two parameter sets in the case of dichotomous outcomes: 1) treatment A and B have equal performance in all treatment periods, i.e. ν_l= 0.5, l=1,...,6; and 2) there is some treatment difference between the two treatments. Let ν_1=0.5, ν_2=0.6, ν_3=0.7, ν_4=0.4, ν_5=0.3, ν_6=0.2. Note that ν_1, ν_2 and ν_3 are the success probabilities of treatment A on period 1, 2

and 3, respectively; while ν_4, ν_5 and ν_6 are the success probabilities for treatment B on period 1, 2 and 3, respectively. Table 4.1 and 4.2 give the estimated number of patients allocated to the treatment sequences for each parameter set. For a simulation with continuous responses, we assume that σ_s^2=2, σ_e^2=1, and μ=100. Two sets of parameters are: 1) π=τ=γ=δ=0 (no treatment difference); and 2) π=τ=γ=2.5, δ=-2.5 (unequal treatment effects). Table 4.3 and 4.4 show the simulation results for the two situations, respectively.

4.4.2 Four-Period Repeated Measurement Design

For p=4, a total of 2^4 (16) treatment sequences are available to compose a design. To simplify the computation, we arbitrarily choose six sequences (ABAA, ABBA, AABA/ BABB, BAAB, BBAB) in simulation because both the design ABAA, ABBA and their duals, and the design ABBA, AABA and their duals are optimal under the self and mixed carryover effects model, as Kunert and Stufken (2008) [10] noted.

As in the two-treatment three-period simulations, we first study simulations with dichotomous outcomes, in which two parameter sets are defined: 1) ν_l= 0.5, l=1,...,8; and 2) ν_1=0.5, ν_2=0.6, ν_3=0.7, ν_4=0.8, ν_5=0.5, ν_6=0.4, ν_7=0.3, and ν_8=0.2, where ν_1, $\ldots \nu_4$ are the success probabilities of treatment A on period 1, 2, 3 and 4, respectively; while ν_5, $\ldots \nu_8$ are the success probabilities for treatment B on period 1, 2, 3, and 4, respectively. Then in simulations with continuous responses, suppose that σ_s^2=2, σ_e^2=1, and μ=100. We also set parameters as follows: 1) π=τ=γ=δ=0; and 2) π=τ=γ=2.5, δ=-2.5.

Table 4.5, 4.6, 4.7, and 4.8 present the four simulation results, respectively.

4.4.3 Six-Period Repeated Measurement Design

For p=6, we consider a case made up of four different treatment sequences. The design consisting of the sequences ABAABA, ABBBAA and their duals is optimal under the self and mixed carryover effects model, according to Kunert and Stufken (2008) [10]. Two parameter sets are similarly defined in both cases of dichotomous and continuous responses: one is in the absence of treatment effects, and the other one is for unequal treatment effects. Table 4.13 and 4.14 summarized the detail of parameter sets for dichotomous, continuous outcomes in various designs. Table 4.9, 4.10, 4.11, and 4.12 correspond to the results from the simulations for six-period designs.

4.4.4 Conclusion

Continuous Responses

Recall that when $\lambda = 0$, the multiple-objective RARMD only concerns itself with the objective of assigning more patients to a better treatment sequences; and when $\lambda = 1$, the design only works for achieving a highly precise estimation rather than taking ethical advantages. For $0 < \lambda < 1$, the RARMD provides a balanced method of performing the two tasks. Therefore, in a design with presence of treatment effect difference and predefined $0 < \lambda < 1$, a treatment sequence with better performance should have more patients assigned, and the differences between the expected numbers of patients of the available sequences achieve most pronounced when λ taking 1.

Table 4.14 gives an example of the expected mean vector $\mathrm{E}(y_{jk})$ for each treatment sequence in various designs when treatment differences are present. For example, in the two-treatment three-period design with a particular pa-

rameter setting, $\mu = 100$, $\pi_2 = \pi_3 = \tau = \gamma = 2.5$ and $\delta = -2.5$, treatment A has a better effect than treatment B because of $\tau = (\tau_A - \tau_B)/2 = 2.5$; $\pi_2 = \pi_3 = 2.5$ indicates that both the second- and third-period effects are 2.5 units higher than the first-period effect; $\gamma = (\gamma_A - \gamma_B)/2 = 2.5$ indicates that self carryover effect of AA increases an overall mean effect by 2.5 units, while the effect of BB reduces the overall mean effect by 2.5 units; similarly, $\delta = (\delta_A - \delta_B)/2 = -2.5$ indicates that the mixed carryover effect of AB decreases an overall mean effect by 2.5 units, while the effect of BA adds an extra 2.5 units to the overall mean effect. If we also assume that a response with higher value represents a better treatment result and that all outcomes are nonnegative, the evaluation function, $f_{j,k}$, can be defined as a monotonic function of the average of the summation of all responses from a specified treatment sequence. Incorporating the given parameter information, we have the expected mean vectors in descending order: AAA, BAA, AAB/ABA/BBB, and ABB/BAB/BAA. Therefore, we expect AAA is the best treatment sequence with most patients assigned into, followed by BAA, then AAB/ABA/BBB, ABB/BAB/BBA last. AAB, ABA, or BBB have similar treatment performance, then their expected numbers of patients allocated would be approximately equal. So do ABB, BAB or BBA. The simulations with four- and six-period RARMDs take similar approaches to estimating the expected mean vectors $\mathrm{E}(y_{jk})$. According to the values of $\mathrm{E}(y_{jk})$ for these sequence groups, shown in Table 4.14, we expect that ABAA and AABA have equal best performances, followed by BAAB, ABBA, and both BABB and BBAB last for p=4; while ABAABA is the best, followed by BAAABB and ABBBAA, and BABBAB last for p=6. The results from Table 4.2, 4.6, and 4.10 confirm such conclusions.

When treatment effects are absent, i.e. π=τ=γ=δ=0, we expect the design will assign an approximately equal number of patients to each treatment sequence if $\lambda = 0$; and the results from Table 4.3, 4.7 and 4.11 are as expected. Meanwhile, if $\lambda = 1$, the design is aimed to maximize the information matrix to improve the estimation precision; and we observe that the sequences which are components of a optimal design under the traditional model with first-order carryover effects stand out from the rest, regardless of whether treatment difference exists or not. Such sequences are ABB/BAA for p=3, ABBA/BAAB for p=4, and ABBBAA/BAAABB for p=6.

Dichotomous Responses

Table 4.1, 4.2, 4.5, 4.6, 4.9, 4.10 present the simulation results in the case of dichotomous responses. The expected outcomes for each treatment sequence in various period designs are summarized in Table 4.13. We use the product of all success probabilities from a given treatment sequence to determine the evaluation function; also, we assume some treatment sequences with larger values of the product will have better treatment performance.

For example, in the designs with p=3, shown in Table 4.13, the values of expected outcomes in descending order is AAA, BAA, ABA, BBA, AAB, BAB, ABB and BBB; for p=4, the descending order is ABAA, AABA, BAAB/ABBA, BBAB, BABB; for p=6, it is ABAABA, BAAABB, ABBBAA, and BABBAB. We expect the estimated numbers of patients in the available treatment sequences in each design therefore would follow such order when the design has unequal treatment effects and λ taking 1 or $0 < \lambda < 1$. The simulation results from Table 4.2, 4.6, 4.10 confirmed such conclusion.

The adaptive designs with dichotomous responses have assignment results

similar to that of continuous responses when $\lambda = 0$: 1) Without treatment difference each treatment sequence will receive approximately equal proportion of patients; and 2) With some treatment difference present, the treatment sequence performing best will have most patients allocated. However, when $\lambda = 1$, the result is different from that for continuous response case: the expected numbers of patients in each sequence are also approximately equal, rather than some sequences standing out; it suggests that these sequences when maximizing the information matrix take the same weights in the case of dichotomous responses. We are unable to make clear explanation for such difference, but suggest that it is partly because the optimal designs we used for dichotomous response here are based on reviewed literature for continuous responses.

N	λ	N_{AAA}	N_{AAB}	N_{ABA}	N_{ABB}	N_{BBB}	N_{BBA}	N_{BAB}	N_{BAA}
12	1	1.328	1.404	1.494	1.761	1.317	1.416	1.508	1.772
12	0.9	1.476	1.519	1.502	1.499	1.491	1.51	1.484	1.519
12	0.7	1.515	1.493	1.491	1.535	1.454	1.517	1.491	1.504
12	0.3	1.51	1.442	1.471	1.522	1.543	1.48	1.516	1.516
12	0	1.406	1.425	1.464	1.441	1.529	1.559	1.578	1.598
40	1	3.847	4.191	5.131	6.831	3.85	4.188	5.128	6.834
40	0.9	4.997	4.969	4.82	5.267	4.752	5.125	4.978	5.092
40	0.7	4.795	4.871	5.273	4.991	5.012	5.042	5.107	4.909
40	0.3	4.963	5.015	5.035	4.64	5.068	5.134	5.278	4.867
40	0	4.954	5.113	5.207	4.945	4.983	5.006	4.824	4.968
80	1	7.633	8.671	10.059	13.637	7.619	8.685	10.073	13.623
80	0.9	10.633	9.671	10.086	9.589	10.487	9.638	10.384	9.512
80	0.7	9.769	10.769	10.348	10.27	8.897	9.482	10.199	10.266
80	0.3	10.128	11.042	9.996	9.953	9.753	9.756	9.464	9.908
80	0	10.107	9.658	9.692	9.506	10.16	9.709	10.813	10.355
100	1	9.636	10.927	12.847	16.59	9.644	10.919	12.839	16.598
100	0.9	11.918	12.614	12.691	12.494	12.224	12.54	12.662	12.857
100	0.7	13.899	12.359	12.27	11.769	12.849	12.89	12.114	11.85
100	0.3	12.042	12.682	12.54	13.421	11.683	12.605	11.67	13.357
100	0	12.603	13.878	10.272	11.696	11.577	12.871	13.316	13.787

Table 4.1: Estimated numbers of subjects for each treatment sequence in the case of dichotomous responses:p=3, no treatment difference

N	λ	N_{AAA}	N_{AAB}	N_{ABA}	N_{ABB}	N_{BBB}	N_{BBA}	N_{BAB}	N_{BAA}
12	1	1.332	1.455	1.449	1.73	1.522	1.391	1.595	1.526
12	0.9	1.478	1.566	1.482	1.497	1.524	1.464	1.539	1.45
12	0.7	1.585	1.511	1.479	1.451	1.416	1.485	1.518	1.555
12	0.3	1.854	1.478	1.587	1.272	1.236	1.476	1.344	1.753
12	0	1.945	1.316	1.533	1.195	1.146	1.439	1.326	2.1
40	1	4.621	4.484	4.876	6.019	4.619	4.487	4.923	5.971
40	0.9	6.862	4.56	5.496	3.778	3.493	4.997	4.617	6.197
40	0.7	9.458	4.11	5.088	2.705	2.42	4.271	3.409	8.539
40	0.3	11.882	3.502	5.577	1.802	1.705	4.023	2.889	8.62
40	0	12.22	3.008	5.718	1.63	1.443	3.906	2.288	9.787
80	1	9.104	8.879	9.973	12.044	9.091	8.892	9.986	12.031
80	0.9	18.208	8.071	10.102	6.273	5.636	9.605	6.816	15.289
80	0.7	25.797	5.835	10.623	3.094	2.822	7.624	4.512	19.693
80	0.3	28.857	4.905	11.551	2.117	1.709	7.614	3.745	19.502
80	0	29.109	4.798	9.74	1.745	1.49	6.757	3.473	22.888
100	1	11.501	11.27	12.339	14.89	11.512	11.259	12.328	14.901
100	0.9	24.505	9.317	12.792	7.07	6.604	11.248	8.084	20.38
100	0.7	33.667	6.512	12.517	3.436	2.761	9.203	4.488	27.416
100	0.3	35.999	6.161	12.643	2.158	1.844	10.017	4.201	26.977
100	0	38.363	5.334	11.807	1.818	1.5	8.305	3.381	29.492

Table 4.2: Estimated numbers of subjects for each treatment sequence in the case of dichotomous responses:p=3, unequal treatment effects

N	λ	N_{AAA}	N_{AAB}	N_{ABA}	N_{ABB}	N_{BBB}	N_{BBA}	N_{BAB}	N_{BAA}
10	1	1	1.007	1.472	1.521	1	1.007	1.487	1.506
10	0.9	1.045	1.165	1.405	1.385	1.032	1.18	1.397	1.391
10	0.7	1.216	1.269	1.27	1.287	1.194	1.204	1.284	1.276
10	0.3	1.221	1.245	1.264	1.258	1.255	1.233	1.251	1.273
10	0	1.226	1.258	1.271	1.264	1.246	1.262	1.22	1.253
40	1	1.005	5.995	5.973	7.027	1.007	5.993	5.974	7.026
40	0.9	4.249	4.877	5.383	5.689	4.222	4.931	5.334	5.315
40	0.7	4.783	5.01	5.047	4.955	4.73	5.241	5.053	5.181
40	0.3	4.9	4.744	5.119	5.009	5.108	4.96	5.07	5.09
40	0	5.099	4.8	4.868	4.938	4.964	5.342	4.841	5.148
80	1	1.008	12.997	11.858	14.137	1.01	12.996	11.854	14.14
80	0.9	9.008	10.523	10.392	10.049	9.634	9.779	10.319	10.296
80	0.7	9.608	10.794	9.969	10.062	9.671	10.257	9.413	10.226
80	0.3	10.13	10.391	10.156	9.819	9.934	9.889	10.048	9.633
80	0	10.201	9.859	10.168	10.047	10.273	9.839	9.86	9.753
100	1	1.006	16.438	14.619	17.937	1.005	16.441	14.631	17.923
100	0.9	11.454	12.594	12.418	12.759	11.678	12.792	13.278	13.027
100	0.7	12.523	12.857	12.454	12.197	11.744	11.921	12.979	13.325
100	0.3	12.019	13.053	12.231	12.114	12.54	12.578	13.03	12.435
100	0	11.984	12.743	12.626	12.546	12.04	13.228	12.779	12.054

Table 4.3: Estimated numbers of subjects for each treatment sequence in the case of continuous responses:p=3, no treatment difference

N	λ	N_{AAA}	N_{AAB}	N_{ABA}	N_{ABB}	N_{BBB}	N_{BBA}	N_{BAB}	N_{BAA}
10	1	1	1.003	1.482	1.515	1	1.003	1.483	1.514
10	0.9	1.167	1.067	1.636	1.142	1.012	1.006	1.111	1.859
10	0.7	1.945	1.106	1.152	1.028	1.09	1.019	1.026	1.634
10	0.3	2.175	1.097	1.108	1.011	1.089	1.012	1.015	1.493
10	0	2.211	1.115	1.094	1.018	1.09	1.01	1.013	1.449
40	1	1.016	5.984	5.964	7.036	1.017	5.983	5.963	7.037
40	0.9	10.293	3.753	5.618	2.727	4.184	2.163	2.687	8.575
40	0.7	13.685	3.817	4.403	2.037	4.022	2.008	2.014	8.014
40	0.3	14.154	4.003	3.992	1.85	3.974	1.853	1.85	8.324
40	0	14.917	3.893	3.757	1.845	3.76	1.802	1.862	8.164
80	1	1.012	12.994	11.854	14.14	1.009	12.997	11.847	14.147
80	0.9	25.222	7.334	9.704	4.191	8.137	3.788	4.025	17.599
80	0.7	29.321	7.672	8.103	3.449	7.824	3.386	3.34	16.905
80	0.3	30.168	7.561	7.936	3.201	7.692	3.288	3.439	16.715
80	0	30.391	7.904	7.507	3.394	7.828	3.035	3.265	16.676
100	1	1.009	16.415	14.627	17.949	1.008	16.415	14.631	17.946
100	0.9	32.989	9.271	11.101	4.997	10.293	4.433	4.993	21.923
100	0.7	37.076	9.706	9.933	3.995	9.667	3.933	4.19	21.5
100	0.3	37.089	9.549	9.944	3.965	9.963	3.958	4.174	21.358
100	0	38.49	9.421	9.333	3.898	9.544	3.943	4.08	21.291

Table 4.4: Estimated numbers of subjects for each treatment sequence in the case of continuous responses:p=3, unequal treatment effects

N	λ	N_{ABAA}	N_{ABBA}	N_{AABA}	N_{BABB}	N_{BAAB}	N_{BBAB}
12	1	1.958	1.822	2.208	1.923	1.871	2.218
12	0.9	2.063	1.909	2.012	2.103	1.903	2.01
12	0.7	2.009	2.036	1.964	2.037	1.96	1.994
12	0.3	1.975	1.955	2.016	1.992	2.041	2.021
12	0	1.943	1.947	1.937	1.983	2.117	2.073
40	1	6.007	6.005	7.988	6.007	6.005	7.988
40	0.9	6.547	6.67	6.764	6.541	6.645	6.833
40	0.7	6.364	6.939	6.583	6.705	6.83	6.579
40	0.3	6.538	6.828	6.796	6.622	6.668	6.548
40	0	5.985	6.409	6.648	6.752	7.161	7.045
80	1	11.979	12.028	15.993	11.979	12.028	15.993
80	0.9	13.373	12.965	13.666	13.245	13.194	13.557
80	0.7	14.038	13.492	12.317	13.889	13.671	12.593
80	0.3	14.269	12.593	13.621	13.925	12.689	12.903
80	0	12.622	13.389	13.349	12.412	13.57	14.658
100	1	14.926	15.037	20.037	14.926	15.037	20.037
100	0.9	16.105	16.532	17.379	15.824	16.646	17.514
100	0.7	15.978	16.638	17.364	16.094	16.712	17.214
100	0.3	15.871	18.063	14.401	17.891	16.865	16.909
100	0	15.525	16.978	16.046	16.89	16.957	17.604

Table 4.5: Estimated numbers of subjects for each treatment sequence in the case of dichotomous responses:p=4, no treatment difference

N	λ	N_{ABAA}	N_{ABBA}	N_{AABA}	N_{BABB}	N_{BAAB}	N_{BBAB}
12	1	2.114	1.764	2.118	2.056	1.791	2.157
12	0.9	2.183	1.842	2.08	2.045	1.881	1.969
12	0.7	2.211	1.839	2.087	1.9	1.992	1.971
12	0.3	2.651	1.902	2.211	1.489	2.021	1.726
12	0	2.764	1.886	2.252	1.377	2.059	1.662
40	1	7.16	5.549	7.335	7.11	5.563	7.283
40	0.9	8.308	5.536	7.589	5.581	7.005	5.981
40	0.7	11.036	5.074	8.015	4.13	6.643	5.102
40	0.3	14.298	5.164	9	2.578	5.652	3.308
40	0	14.496	5.461	9.041	2.125	5.745	3.132
80	1	13.832	11.36	14.808	13.832	11.36	14.808
80	0.9	20.293	10.103	15.346	9.781	14.466	10.011
80	0.7	29.602	9.218	16.155	5.605	12.401	7.019
80	0.3	33.789	9.462	18.839	3.039	10.187	4.684
80	0	35.203	8.56	20.139	2.813	8.499	4.786
100	1	17.278	14.203	18.521	17.275	14.204	18.519
100	0.9	27.096	11.576	20.133	11.251	17.635	12.309
100	0.7	38.128	10.083	22.809	5.831	14.832	8.317
100	0.3	46.182	10.395	24.85	2.707	11.189	4.677
100	0	46.044	8.827	26.297	2.376	12.263	4.193

Table 4.6: Estimated numbers of subjects for each treatment sequence in the case of dichotomous responses:p=4, unequal treatment effects

N	λ	N_{ABAA}	N_{ABBA}	N_{AABA}	N_{BABB}	N_{BAAB}	N_{BBAB}
10	1	1.208	2.383	1.409	1.208	2.383	1.409
10	0.9	1.627	1.731	1.634	1.659	1.706	1.643
10	0.7	1.707	1.624	1.68	1.67	1.689	1.63
10	0.3	1.675	1.648	1.618	1.732	1.643	1.684
10	0	1.618	1.675	1.656	1.716	1.652	1.683
40	1	3.668	10	6.332	3.668	10	6.332
40	0.9	6.727	6.667	6.618	6.518	6.999	6.471
40	0.7	6.697	6.298	6.694	6.823	6.739	6.749
40	0.3	6.777	6.51	6.674	6.852	6.582	6.605
40	0	7.111	6.612	6.465	6.861	6.439	6.512
80	1	6.966	20	13.034	6.966	20	13.034
80	0.9	13.136	13.183	13.669	12.187	14.251	13.574
80	0.7	13.897	13.082	13.55	12.129	13.666	13.676
80	0.3	13.408	13.587	13.663	12.383	13.694	13.265
80	0	12.259	13.09	13.565	13.579	14.261	13.246
100	1	8.617	25	16.383	8.617	25	16.383
100	0.9	16.348	17.564	16.411	16.578	16.799	16.3
100	0.7	16.434	17.272	17.138	16.173	16.31	16.673
100	0.3	16.24	16.441	17.315	16.822	17.439	15.743
100	0	16.83	16.032	16.783	16.455	16.628	17.272

Table 4.7: Estimated numbers of subjects for each treatment sequence in the case of continuous responses:p=4, no treatment difference

N	λ	N_{ABAA}	N_{ABBA}	N_{AABA}	N_{BABB}	N_{BAAB}	N_{BBAB}
10	1	1.232	2.354	1.414	1.232	2.354	1.414
10	0.9	2.132	1.207	2.105	1.294	1.995	1.267
10	0.7	2.386	1.19	2.364	1.123	1.807	1.13
10	0.3	2.49	1.226	2.47	1.093	1.615	1.106
10	0	2.463	1.221	2.552	1.071	1.607	1.086
40	1	3.668	10	6.332	3.668	10	6.332
40	0.9	10.888	3.114	11.001	3.229	8.558	3.21
40	0.7	11.575	3.588	12.048	2.666	7.51	2.613
40	0.3	12.085	3.459	12.353	2.406	7.225	2.472
40	0	12.607	3.533	12.457	2.394	6.688	2.321
80	1	6.952	20	13.048	6.952	20	13.048
80	0.9	23.051	6.519	23.273	5.063	16.535	5.559
80	0.7	24.438	6.215	25.581	4.574	14.728	4.464
80	0.3	24.197	6.842	25.325	4.662	14.662	4.312
80	0	26.176	6.401	24.852	4.303	13.852	4.416
100	1	8.633	25	16.367	8.633	25	16.367
100	0.9	29.387	7.981	29.875	6.491	20.081	6.185
100	0.7	30.454	8.314	31.945	5.241	18.394	5.652
100	0.3	31.585	8.543	31.792	5.535	16.802	5.743
100	0	32.607	8.224	31.487	5.226	17.011	5.445

Table 4.8: Estimated numbers of subjects for each treatment sequence in the case of continuous responses:p=4, unequal treatment effects

N	λ	N_{ABAABA}	N_{ABBBAA}	N_{BABBAB}	N_{BAAABB}
12	1	2.948	3.073	2.924	3.055
12	0.9	3.026	2.961	3.032	2.981
12	0.7	3.037	2.942	3.021	3
12	0.3	2.897	3.069	2.947	3.087
12	0	2.928	3.024	2.991	3.057
40	1	9.603	10.397	9.603	10.397
40	0.9	9.921	10.106	9.925	10.048
40	0.7	10.221	9.831	10.14	9.808
40	0.3	10.2	9.539	10.712	9.549
40	0	9.988	10.16	10.067	9.785
80	1	19.344	20.656	19.344	20.656
80	0.9	20.322	19.733	20.284	19.661
80	0.7	19.829	20.348	20.078	19.745
80	0.3	18.978	19.703	20.72	20.599
80	0	19.459	20.531	18.966	21.044
100	1	24.145	25.855	24.145	25.855
100	0.9	25.343	24.578	25.411	24.668
100	0.7	24.545	25.698	24.777	24.98
100	0.3	25.126	25.7	24.661	24.513
100	0	24.472	24.493	23.918	27.117

Table 4.9: Estimated numbers of subjects for each treatment sequence in the case of dichotomous responses:p=6, no treatment difference

N	λ	N_{ABAABA}	N_{ABBBAA}	N_{BABBAB}	N_{BAAABB}
12	1	2.902	3.201	3.025	2.872
12	0.9	2.983	3.212	2.97	2.835
12	0.7	2.924	3.179	3.079	2.818
12	0.3	2.993	3.553	2.964	2.49
12	0	3.017	4	2.868	2.115
40	1	9.814	10.171	9.835	10.18
40	0.9	9.951	10.519	9.88	9.65
40	0.7	10.359	11.462	10.018	8.161
40	0.3	9.551	16.137	8.575	5.737
40	0	9.325	16.024	9.446	5.205
80	1	19.611	20.389	19.611	20.389
80	0.9	19.671	22.223	19.162	18.944
80	0.7	20.201	26.47	18.406	14.923
80	0.3	18.141	39.469	15.137	7.253
80	0	21.764	35.313	14.757	8.166
100	1	24.457	25.543	24.457	25.543
100	0.9	24.853	28.212	23.967	22.968
100	0.7	24.243	35.309	22.313	18.135
100	0.3	24.716	50.929	16.135	8.22
100	0	26.732	44.863	19.196	9.209

Table 4.10: Estimated numbers of subjects for each treatment sequence in the case of dichotomous responses:p=6, unequal treatment effects

N	λ	N_{ABAABA}	N_{ABBBAA}	N_{BABBAB}	N_{BAAABB}
10	1	1	4	1	4
10	0.9	1.99	2.999	1.982	3.029
10	0.7	2.422	2.544	2.456	2.578
10	0.3	2.446	2.505	2.512	2.537
10	0	2.548	2.493	2.476	2.483
40	1	5	15	5	15
40	0.9	9.601	10.595	9.287	10.517
40	0.7	9.289	10.285	9.87	10.556
40	0.3	10.086	9.484	9.869	10.561
40	0	9.936	10.078	10.08	9.906
80	1	9	31	9	31
80	0.9	19.874	20.629	19.954	19.543
80	0.7	20.027	20.257	19.281	20.435
80	0.3	19.555	20.35	19.834	20.261
80	0	19.08	20.783	20.919	19.218
100	1	12	38	12	38
100	0.9	25.377	24.921	23.846	25.856
100	0.7	23.188	26.153	24.881	25.778
100	0.3	25.88	25.146	24.968	24.006
100	0	24.064	26.18	24.438	25.318

Table 4.11: Estimated numbers of subjects for each treatment sequence in the case of continuous responses:p=6, no treatment difference

N	λ	N_{ABAABA}	N_{ABBBAA}	N_{BABBAB}	N_{BAAABB}
10	1	1	4	1	4
10	0.9	2.357	2.753	1.774	3.116
10	0.7	3.471	2.096	1.653	2.78
10	0.3	3.752	2	1.468	2.78
10	0	3.848	1.948	1.493	2.711
40	1	5	15	5	15
40	0.9	13.858	8.611	6.287	11.244
40	0.7	16.073	7.32	5.027	11.58
40	0.3	16.838	7.255	4.744	11.163
40	0	16.63	7.306	4.818	11.246
80	1	9	31	9	31
80	0.9	30.034	15.625	10.479	23.862
80	0.7	33.541	15.096	9.034	22.329
80	0.3	35.176	13.643	9.536	21.645
80	0	34.812	14.127	8.657	22.404
100	1	12	38	12	38
100	0.9	38.609	19.641	13.332	28.418
100	0.7	43.851	16.706	11.522	27.921
100	0.3	42.006	17.201	11.243	29.55
100	0	44.521	17.299	10.423	27.757

Table 4.12: Estimated numbers of subjects for each treatment sequence in the case of continuous responses:p=6, unequal treatment effects

Design	Parameter	Treatment sequence	Expected responses
Two-treatment three-period	$\nu_{A1} = 0.5$,$\nu_{B1} = 0.4$	AAA	$(0.5, 0.6, 0.7)'$
	$\nu_{A2} = 0.6$,$\nu_{B2} = 0.3$	AAB	$(0.5, 0.6, 0.2)'$
	$\nu_{A3} = 0.7$,$\nu_{B1} = 0.2$	ABA	$(0.5, 0.3, 0.7)'$
		ABB	$(0.5, 0.3, 0.2)'$
		BBB	$(0.4, 0.3, 0.2)'$
		BBA	$(0.4, 0.3, 0.7)'$
		BAB	$(0.4, 0.6, 0.2)'$
		BAA	$(0.4, 0.6, 0.7)'$
Two-treatment four-period	$\nu_{A1} = 0.5$,$\nu_{B1} = 0.5$	ABAA	$(0.5, 0.4, 0.7, 0.8)'$
	$\nu_{A2} = 0.6$,$\nu_{B2} = 0.4$	ABBA	$(0.5, 0.4, 0.3, 0.8)'$
	$\nu_{A3} = 0.7$,$\nu_{B3} = 0.3$	AABA	$(0.5, 0.6, 0.3, 0.8)'$
	$\nu_{A4} = 0.8$,$\nu_{B4} = 0.2$	BABB	$(0.5, 0.6, 0.3, 0.2)'$
		BAAB	$(0.5, 0.6, 0.7, 0.2)'$
		BBAB	$(0.5, 0.4, 0.7, 0.2)'$
Two-treatment six-period	$\nu_{A1} = 0.40$,$\nu_{B1} = 0.80$	ABAABA	$(0.4, 0.75, 0.6, 0.65, 0.4, 0.8)'$
	$\nu_{A2} = 0.50$,$\nu_{B2} = 0.75$	ABBBAA	$(0.4, 0.75, 0.65, 0.5, 0.75, 0.8)'$
	$\nu_{A3} = 0.60$,$\nu_{B3} = 0.65$	BABBAB	$(0.8, 0.5, 0.65, 0.5, 0.75, 0.3)'$
	$\nu_{A4} = 0.65$,$\nu_{B4} = 0.50$	BAAABB	$(0.8, 0.5, 0.6, 0.65, 0.4, 0.3)'$
	$\nu_{A5} = 0.75$,$\nu_{B5} = 0.40$		
	$\nu_{A6} = 0.80$,$\nu_{B6} = 0.30$		

Table 4.13: Expected response for each treatment sequence in the case of unequal treatment effects with dichotomous outcomes

Design	Parameter Set	Treatment sequence	Expected responses
Two-treatment three-period	$\mu = 100$	AAA	$(102.5, 107.5, 107.5)'$
	$\pi_2 = \pi_3 = \tau = \gamma = 2.5$	AAB	$(102.5, 107.5, 97.5)'$
	$\delta = -2.5$	ABA	$(102.5, 97.5, 107.5)'$
		ABB	$(102.5, 97.5, 97.5)'$
		BBB	$(97.5, 97.5, 97.5)'$
		BBA	$(97.5, 97.5, 107.5)'$
		BAB	$(97.5, 107.5, 97.5)'$
		BAA	$(97.5, 107.5, 107.5)'$
Two-treatment four-period	$\mu = 100$	ABAA	$(102.5, 97.5, 107.5, 107.5)'$
	$\pi_2 = \pi_3 = \pi_4 = 2.5$	ABBA	$(102.5, 97.5, 97.5, 107.5)'$
	$\tau = \gamma = 2.5$	AABA	$(102.5, 107.5, 97.5, 107.5)'$
	$\delta = -2.5$	BABB	$(97.5, 107.5, 97.5, 97.5)'$
		BAAB	$(97.5, 107.5, 107.5, 97.5)'$
		BBAB	$(97.5, 97.5, 107.5, 97.5)'$
Two-treatment six-period	$\mu = 100$	ABAABA	$(102.5, 97.5, 107.5, 107.5, 97.5, 107.5)'$
	$\pi_2 = \ldots = \pi_6 = 2.5$	ABBBAA	$(102.5, 97.5, 97.5, 97.5, 107.5, 107.5)'$
	$\tau = \gamma = 2.5$	BABBAB	$(97.5, 107.5, 97.5, 97.5, 107.5, 97.5)'$
	$\delta = -2.5$	BAAABB	$(97.5, 107.5, 107.5, 107.5, 97.5, 97.5)'$

Table 4.14: Expected response for each treatment sequence in the case of unequal treatment effects with continuous outcomes

Chapter 5

Data Analysis

In this chapter, we use some published N-of-1 trials data under the traditional linear mixed model and self and mixed carryover effect model, respectively, to compare the efficiency in estimating treatment effect.

Introduction

The following data were from a series of 58 N-of-1 trials that compare amitriptyline(AMT) therapy and the combination of AMT and fluoxetine(FL) for treating fibromyalgia syndrome. Each N-of-1 trial had six treatment periods with three sets of treatment pair. For each paired treatments, one period is on AMT, another is on AMT+FL. The treatment assignment was block randomized in the first pair, and simply randomized with the start medication in the other two pairs. The Fibromyalgia Impact Questionnaire (FIQ) score, as the main interest responses variable, was collected once at the end of each six 1-week treatment periods. In the following analysis we only used the data from the 46 patients who completed at least two treatment periods.

The Models

To investigate the efficiency for estimation of the treatment effect under different models, we reanalyze the data from this series of N-of-1 trials in various scenarios. For example, if we use the results only from the first period, the analysis would mirror a completely randomized trial. Or if we analyze the response from the first three periods with their treatment allocation information, this would correspond to a standard three-period crossover design.

For these N-of-1 trials, there are total eight different treatment sequences available. Due to the characteristics of randomization order, we can mirror the standard CODs with at most three treatment periods and RCT. Therefore, two three-period CODs are considered : 1) the design ABB/BAA, denoted as COD(2,3,2), where (2,3,2) means two treatments, three periods, and two sequences. We chose this design because its dual balanced design is optimal over the class of three-period designs under the traditional linear mixed model; 2) the design ABB,ABA/BAA,BAB, denoted as COD(2,3,4). Since it is one of optimal designs for estimation the treatment difference (τ) under the self and mixed carryover effect model, and the four treatment sequences are available to group from the N-of-1 trials, we include this design in our analysis.

Suppose that all models under investigation have random subject effects with an equicorrelated covariance structure. First, recall that the traditional linear mixed model, which is defined as Model I in this chapter, is given in Equation 3.1,

$$y_{ijk} = \mu + \pi_i + \mathrm{I}_{d(i,k)}\tau + \mathrm{I}_{d(i-1,k)}\gamma + \zeta_{jk} + \epsilon_{ijk} \qquad \text{(I)}$$

here we assume that each treatment has a carryover effect which is the same regardless of the treatment in the next period.

An alternative of the traditional model for the crossover designs with carryover effect, is called the self and mixed carryover effect model. Throughout this chapter we define this model as Model II. Instead of assuming that the carryover effect of a treatment does not interact with the direct effect of the treatment in the following period, Model II assumes that each treatment has two different carryover effects for every treatment in the next period. Model II is introduced in chapter 4,

$$y_{ijk} = \mu + \pi_i + \tau_{d_k(i,j)} + \mathrm{I}_{ijk}\gamma_{d_k(i-1,j)} + (1 - \mathrm{I}_{ijk})\delta_{d_k(i-1,j)} + \zeta_{jk} + \epsilon_{ijk} \qquad \text{(II)}$$

Because we assume that the carryover effect of a treatment on itself is different from the carryover effect on other treatments, we call it as the self carryover effect in the former case, and mixed carryover effect in the latter case.

Finally, when the assumptions of no period and carryover effects are valid, we consider a simple linear mixed model as Model III, in which the response only depends on the treatment effect and random subject effects. Model III is also reviewed in chapter 2,

$$y_j = \mu + \zeta_j + \epsilon_j \qquad \text{(III)}$$

Results

Table 5.1 presents results of the analysis that mirror RCT and CODs under the traditional linear mixed model (I), self and mixed effect model (II), and

Design	Model								
	I			II			III		
	τ^a	(SE)	P value	τ	(SE)	P value	τ	(SE)	P value
RCT	-5.25	(4.59)	0.2590						
ABB/BAA[b]	-1.64	(3.13)	0.6022	1.72	(5.24)	0.7435	-2.42	(3.30)	0.4638
ABB/BAA, ABA/BAB[c]	-4.06	(2.77)	0.1456	-4.80	(4.60)	0.2981	-6.37	(2.65)	0.0184
N-of-1 trials	-4.95	(2.12)	0.0203	-4.07	(4.58)	0.3749	-6.07	(1.84)	0.0011

[a] τ=FIQ score on AMT+FL - FIQ score on AMT
[b] The number of patients in this design, $N_{COD(2,3,2)}$=22
[c] The number of patients in this design, $N_{COD(2,3,4)}$=41

Table 5.1: Mean treatment effect difference

Model (III) omitting period and carryover effects. For the analysis of RCT and N-of-1 trials under Model III, the same results are reported by Zucker (2010). In a two-treatment three-period repeated measurement design based on Model I, for estimation of τ, the design ABB/BAA with an equal number of subjects in each sequence is optimal. However, in our example, the total subjects in the design ABB/BAA is $N_{COD(2,3,2)}$=22 and $N_{COD(2,3,4)}$=41 in the design ABB,ABA/BAA,BAB, thus we did not see an expected smaller standard error of τ in the design ABB/BAA.

These results are summarized below:

1. In the COD(2,3,2) with ABB/BAA, the carryover effect is marginal significant in Model I with P-value 0.0405 (not shown).

2. For the COD(2,3,4) with ABB,ABA/BAA,BAB, the carryover effect is significant with P-value 0.0149 (not shown) in Model I.

3. Under Model II, both COD(2,3,2) and COD(2,3,4) do not find the carryover or periods effects significant.

4. The N-of-1 trials design under both model I and II, has not significant period or carryover effect terms.

5. Based on the Model (III), both COD(2,3,4) and N-of-1 trials design have found that the treatment effect is significant.

Under Model(II), all results lead to the same conclusion: the treatment effect is not significant and all the estimates of τ have almost the same standard errors. The findings suggest that there is little to be gained by using N-of-1 trials design with the self and mixed carryover effect model when it is compared to RCT and COD(2,3,4) with sequences ABB,ABA/BAA,and BAB.

Mean Responses Analysis

Recall that Model III in matrix form is given by Equation (2.1),

$$y_j = X_j\beta + Z_j b_j + \varepsilon_j$$

and β is a vector of fixed effects that are assumed to be the same for all subjects; $X_j\beta$ thus represents an overall mean. The b_j is the subject-specific effect vector and has a normal distribution. Combining these two terms, the mean response for the jth subject is

$$\mathrm{E}(y_j|b_j) = X_j\beta + Z_j b_j$$

A same patient's results from the N-of-1 trial and the COD(2,3,4) are collected in Appendix A. An illustration of the mean responses analysis for this patient is then presented in Table 5.2. Model III has intercepts that vary

Patient 101	Responses y_j			Predicted Responses E($y_j|b_j$)	
				N-of-1 trial	COD(2,3,4)
trt 1	29.28571,	39.33036,	29.25595	41.95	46.34
trt 2	51.86012,	11.9494,	57.38095	35.88	39.97

Table 5.2: A single patient's response profile

randomly among subjects; it also allows the mean value of the intercept to differ in the two treatment groups.

Chapter 6

Conclusions

This thesis provides a general review on some topics of interest in clinical trial designs. In the first part, we provided the basic introduction to the common clinical trials, and focused our discussion on the repeated measurement designs such as crossover trials and N-of-1 trials.

Crossover trials are appealing in clinical trials as they allow evaluation of each patient's outcomes on all of the treatments under investigation. Since each patient is used as its own control, variability is reduced because the within-subject difference is usually smaller than the between-subject difference. This reduction in variability enables the clinical researchers and investigators to use a relatively small number of patients to detect treatment differences. To appreciate its increased efficiency, we compared some optimal crossover trials with three, four and six periods to the completely randomized design and N-of-1 trial design using different evaluation criteria, such as the relative design efficiency and cost efficiency comparison. In brief, the three-period optimal crossover trials performed very competitively with various model assumptions. Three linear mixed models used for these designs

are also under examination: 1) the tradition model with first order carryover effects; 2) the self and mixed carryover effects model; 3) simple linear mixed model omitting carryover effects.

The combined N-of-1 trial design is a special case of multi-crossover design but it emphasizes the single-patient assessment. It borrows as much strength as it can from a usual crossover trial and possesses the advantages of the single patient trial design. In an N-of-1 trial, the treatment tests are patient friendly and easy to implement. Individual patient responses are used to guide the patient-physician treatment decisions and each patient personally benefits from evidence based prescribing. The efficacy, side effect, and other individual-specific characteristics data from the trial can be combined to provide information with regard to the efficiency of treatments for an individual patient. Much attention has been given to the N-of-1 design because of its individual benefits and ethical advantages. However, current methods in use need improvement to fully interpret N-of-1 trial data. In this thesis, we only reviewed the linear mixed model analysis regarding N-of-1 trials.

In general, additional data increase efficiency. The example in Chapter Five shows that the N-of-1 trial design increased efficiency of estimation compared with the RCT and the COD with ABB, ABA/BAA, BAB under the simple linear mixed model. However, a question raised relates to the trade-off between increasing the periods for a more precise estimation if choosing N-of-1, and increasing the cost efficiency if choosing the COD (2, 3, 4). Under the usual linear mixed model with carryover effects, an N-of-1 design gains little statistical power to test treatment differences, even with increased treatment periods.

As we discussed before, each design has its particular advantages and dis-

advantages when applied in clinical setting. The most attractive offer of the COD is its competitive efficiency over the RCT, but with the concern of carryover effects. N-of-1 trials give priority to individual patient evaluation or treatment, making individual estimates of treatment effect as their primary purpose. When we analyze data from N-of-1 trials by using the linear mixed model, we develop the estimates of the overall mean treatment effect of a population, but fail to appreciate its merits of being individual trials, therefore, such design efficiency or cost saving comparison for N-of-1 trials may lead to less important consideration in clinical investigation.

Another way of comparing these clinical trial designs to each other is from the view of randomization application. The RCT, COD and N-of-1 trial are standard randomized designs, which usually employ a simple randomization allocating subjects to each treatment group equally. To achieve a high statistical power for the comparison of treatment effects, this is often the best approach. However, the equal treatment allocation may pose ethical infeasibility or concerns in some clinical settings. To cope with such situations, alternative designs such as response-adaptive designs have been advocated. Such designs can adjust the treatment allocation rule based on patient responses already accrued in the trial, and assign more patients to the better treatment groups.

In that respect, a multiple-objective RARMD for clinical trials has been reviewed in the last part of the thesis. We discussed the adaptive treatment allocation scheme and the statistical models used for such designs. A simulation study presented an illustration of the use of the adaptive treatment allocation scheme.

To incorporate a variety of patient-specific characteristic information and strengthen its advantage as a single patient trial, the N-of -1 trial design re-

quires more complicated analytical approaches. One possible way to better appreciate the attractive properties is to use the Bayesian hierarchical method. Therefore studying this method or searching for an optimal solution to the N-of-1 trial will be the focus of my future research.

Though methodological research on response-adaptive designs for clinical trials has developed rapidly, they have rarely been used in practice. According to Simon (1977)[18], adaptive treatment allocation methods have some inherit limitations, and most published methods have important deficiencies; both reasons render the adaptive design difficult in application. Such limitations or defects, for example, could be possible delays in observing or reporting responses, and difficult to make appropriate evaluation or decision rules to choose good treatment sequences. Exploring in depth such topics in the field of clinical trials also interest me as there is much room for improvement of current methods.

Bibliography

[1] P. Armitage. *Sequential Medical trials.* Blackwell:Oxford, 1975.

[2] B.W., Jr. Brown. The crossover experiment for clinical trials. *Biometrics*, 36:69–79, 1980.

[3] D.P. Byar, R.M. Simon, W.T. Friedewald, J.J. Schlesselman, D.L. DeMets, J.H. Ellenberg, M.H. Gail, and J.H. Ware. Randomized clincal trials, perspectives on some recent ideas. *The New England journal of medicine*, 295(2):74–80, 1976.

[4] K.C. Carriere. Crossover designs for clinical trials. *Statistics in Medicine*, 13:1063–1069, 1994.

[5] K.C. Carriere and R. Huang. Crossover designs for two-treatment clinical trials. *Journal of Statistical Planning and Inference*, 87:125–134, 2000.

[6] K.C. Carriere and Y. Liang. Response-adaptive repeated measurement designs for clinical trials (In press). *Biometrics.*

[7] K.C. Carriere and G.C. Reinsel. Investigation of dual-balanced crossover designs for two treatments. *Biometrics*, 48:1157–1164, 1992.

[8] W. G. Cochran and G. M. Cox. *Experimental designs. 2nd ed.* Wiley, New York, 1957.

[9] J.E. Grizzle. The two-period change-over design and its use in clinical trials. *Biometrics*, 21:467–480, 1965.

[10] J. Kunert and J. Stufken. Optimal crossover designs for two treatments in the presence of mixed and self-carryover effects. *Journal of the American Statistical Association*, 103:1641–1647, 2008.

[11] H.B. Kusher. Allocation rules for adaptive repeated measurements designs. *Journal of Statistical Planning and Inference*, 113:293–313, 2003.

[12] E. Laska, M. Meisner, and H.B. Kushner. Optimal crossover designs in the presence of carryover effects. *Biometrics*, 39:1087–1091, 1983.

[13] E. M. Laska and M. Meisner. A variational approach to optimal two-treatment crossover designs: application to carryover-effect models. *Journal of the American Statistical Association*, 80:704–710, 1985.

[14] Y. Liang. *Response-adaptive repeated measurement designs for clinical trials*. PhD. Thesis. University of Alberta, 2006.

[15] Y. Liang and K.C. Carriere. Multiple-objective response-adaptive repeated measurement designs for clinical trials. *Journal of Statistical Planning and Inference*, 139:1134–1145, 2009.

[16] S. Piantadosi. *Clinical trials, a methodological perspective. 2nd ed.* John Wiley & Sons, Inc., Hoboken, New Jersey, 2005.

[17] R. Simon. The randomized play-the-winner rule in medical trials. *Biometrics*, 33:743–749, 1977.

[18] L.J. Wei and S. Durham. Adaptive treatment assignment methods and clinical trials. *Journal of the American Statistical Association*, 73:840–843, 1978.

[19] M. Zelen. Play the winner rule and the controlled clinical trial. *Journal of Clincal Epidemiology*, 64:131–146, 1969.

[20] D.R. Zucker, R. Ruthazer, and C.H. Schmid. Individual (N-of-1)trials can be combined to give population comparative treatment effect estimates:methodologic considerations. *Journal of Clincal Epidemiology*, 63:1312–1323, 2010.

[21] D.R. Zucker, C.H. Schmid, M.W. Mclntosh, R.B. D'Agostino, H.P. Selker, and J. Lau. Combining single patient (N-of-1) trials to estimate population treatment effects and to evaluate individual patient responses to treatment. *Journal of Clincal Epidemiology*, 50:401–410, 1997.

Appendix: SAS outputs for the example in Chapter 5

1. The results for a same patient from N-of-1 trial and crossover design, respectively.

N-of-1 Trial						
Solution for Fixed Effects						
Effect	trt	Estimate	Standard Error	DF	tValue	Pr > |t|
Intercept		42.2547	1.928	0	21.92	.
trt	1	6.0729[a]	1.8387	193	3.3	0.0011
trt	2	0	.	.	.	.
Solution for Random Effects						
Effect	patid	Estimate	Std Err Pred	DF	tValue	Pr > |t|
Intercept	101	-6.3509	5.1089	193	-1.24	0.2153
patid	101	-0.02273	0.5635	193	-0.04	0.9679

Crossover design with ABB,ABA/BAA,BAB						
Solution for Fixed Effects						
Effect	trt	Estimate	Standard Error	DF	tValue	Pr > |t|
Intercept		41.5703	2.0738	0	20.05	.
trt	1	6.3703[b]	2.6513	86	2.4	0.0184
trt	2	0	.	.	.	.
Solution for Random Effects						
Effect	patid	Estimate	Std Err Pred	DF	tValue	Pr> |t|
Intercept	101	-1.5988	4.6401	86	-0.34	0.7313
patid	101	-0.00003	0.02508	86	0	0.9989

[a] trt1=AMT, trt2=AMT+FL, and τ= trt2-trt1.

[b] See a.