Response-Adaptive Repeated Measurement Designs for Clinical Trials

Yuanyuan Liang[1] , Kimmie Chough Carriere[2]

[1]Alberta Research Centre for Child Health Evidence, Department of Pediatrics, University of Alberta
[2]Department of Mathematical and Statistical Sciences, University of Alberta

Abstract

In a response-adaptive design (RAD), we review and update the trial on the basis of outcomes in order to achieve a specific goal. Optimal designs for clinical trials are usually constructed under a single objective. In this paper, we develop a new adaptive allocation rule to improve the current strategies of building response adaptive designs to construct multiple-objective designs. The purpose of this new rule is to increase both estimation precision and treatment benefits by assigning more patients to a better treatment. We demonstrate that the designs constructed under the new proposed allocation rule are more efficient than fixed optimal designs in terms of both the mean squared error and improved patient care.

Keywords: response-adaptive design (RAD), multiple-objective designs, self and simple mixed carryover effects model, evaluation function, optimality criteria.

1. Introduction

There has been a growing interest in developing clinical trials that are designed to ensure that the treatment allocation strategy is better informed from all available sources. This interest has been fueled, in part, by evidence that some health interventions are largely ineffective and even harmful, and therefore a waste of public resources and unethical. In *response-adaptive designs* (RAD), we modify the trial on the basis of outcomes in order to achieve a specific goal. Optimal designs have traditionally been constructed to satisfy a single objective. For example, Kushner (2003) proposed an adaptive allocation rule that was developed by replacing the classical optimal design strategy with a response adaptive allocation method. The purpose was to maximize the information matrix for treatment effects based on already-observed subjects. In classical sequential trials, the decision to terminate the trial reflects a concern that sample size will be minimal (Armitage, 1975). In play-the-winner designs, the goal is to maximize the number of subjects receiving a better treatment (Zelen, 1969). However, such single-objective designs leave something to be desired since investigators often need to deal with more than one objective when designing an experiment (Moerbeek and Wong, 2002).

In this paper, we aim to construct efficient multiple-objective designs that will both increase estimation precision and improve patient care by maximizing the proportion of patients receiving a better treatment. Since constructing a design is highly sensitive to the model employed, we use a model that accommodates several types of carryover effects and random subject effects to discuss the robustness of designs using other models. We also discuss applications to experiments using continuous responses.

Section 2 describes the general model we consider in this paper. Section 3 defines an evaluation function and the optimal design selection criteria for measuring the performance of a treatment sequence and incorporating multiple objectives. We then present the new allocation rule in Section 4. Adaptive designs are constructed in Section 5, followed by conclusions and suggestions for further work.

2. The Model

Traditionally, models for RMD have accommodated simple first-order carryover effects (for example Hedayat and Afsarinejad, 1978). Carriere (1994b) considered some alternate models, limiting the study to three period designs. Afsarinejad and Hedayat (2002) proposed a model that allows for two different types of carryover effects from each treatment—the self and mixed carryover effects. They also studied optimal two-period repeated measurements designs with two or more treatments using a model with fixed subject effects.

In this paper, we consider a model that accommodates random subject effects along with self and mixed carryover effects as

$$y_{ijk} = \mu + \pi_i + \tau_{d_k[i,j]} + (1-\delta_{ijk})\gamma_{d_k[i-1,j]} + \delta_{ijk}\varphi_{d_k[i-1,j]} + \xi_{jk} + \varepsilon_{ijk} \qquad (1)$$

where y_{ijk} denotes the response variable for the j^{th} subject given that treatment sequence k in period i, μ is an overall mean, π_i and ξ_{jk} are the period and subject effects, respectively, $d_k[i, j]$ denotes the treatment used for subject j given treatment sequence k in period i, $i = 1, 2, ..., p$, $j = 1, 2, ..., N_k$, $k = 1, 2, ..., s$, N_k is the number of subjects in sequence k, s is the total number of treatment sequences, and $N = \sum_k N_k$ is the total number of subjects in the study. Both $\gamma_{d_k[i-1,j]}$ and $\varphi_{d_k[i-1,j]}$ represent carryover effects, while δ_{ijk} is an indicator variable, taking 1 if $d_k[i, j] = d_k[i-1, j]$ and 0 otherwise. Thus $\gamma_{d_k[i-1,j]}$ is the carryover effect of one treatment followed by a different treatment, called the

mixed carryover effect, while $\varphi_{d_k[i-1,j]}$ is the carryover effect from a treatment onto itself, called *self carryover effect*, with $\gamma_{d_k[0,j]} = \varphi_{d_k[0,j]} = 0$. ξ_{jk} and ε_{ijk} are mutually independent random effects with mean 0 and variance σ_ξ^2 and σ_ε^2, respectively.

3. Evaluation Function and the Selection Criteria

3.1 Evaluation Function

Performance-driven and multi-objective designs are commonly used in engineering research. For example, one may need to simultaneously optimize various performance measures when designing a vehicle suspension. These performance measures capture desirable properties such as ride comfort and handling. Here, we introduce the concept of performance evaluation to repeated measurement designs. In order to assign more patients to better-performing treatment sequences, we need a mechanism to quantitatively evaluate the performance of each treatment sequence. Therefore, we will define an *evaluation function* for a treatment sequence.

Properties of an Evaluation Function: An evaluation function, $g_k(\cdot)$, for treatment sequence k based on the existing data satisfies the following properties:
1) It is non-negative;
2) It is monotonic.

Obviously, defining an evaluation function for a given treatment sequence is not unique. As long as the above two properties are satisfied, one can define various types of evaluation functions. Based on a pre-defined evaluation function, when two treatment sequences have the same values, we say that the performance of these two treatment sequences is *indistinguishable*.

The criterion issue, that is, the explicit definition of treatment success, is a pervasive problem. In the context of clinical trials for medicinal products, we can define an evaluation function on the basis of the desirable characteristics of the primary outcome. Following are two examples.

Example 1: Venipuncture and intravenous (IV) line placement is a procedure that is commonly performed in pediatric emergency departments. The first attempt at line placement is often unsuccessful, and further attempts follow. To investigate whether using ultrasound instead of the traditional landmark technique can improve the success rate of peripheral IV placement, one can consider a synthetic, two-treatment two-period repeated measurement design, where 10 patients are randomly assigned to the ultrasound technique (A) or the conventional landmark technique (B), and dichotomous responses, first attempt success rate (1 if success, 0 if failure) are collected. The binary data are given in Table 1.

We advocate the idea from the play-the-winner rule and evaluate the performance of a treatment sequence by calculating the probability of success over all subjects, given that particular treatment sequence. Thus, an evaluation function for treatment sequence k can be defined as $g_k = \|S_K\| / \|K\|$, where $\|S_K\|$ denotes the total number of successes for patients given treatment sequence k in all periods, and $\|K\|$ denotes the total number of patients given treatment sequence k.

In our example, the corresponding value for each possible treatment sequence, based on the above evaluation function becomes: $g_{AA} = 5/3 = 1.67$; $g_{AB} = 2/2 = 1$; $g_{BA} = 2/3 = 0.67$; and $g_{BB} = 2/2 = 1$. Therefore, the observed data indicates that treatment sequence AA is the best of these four possible treatment sequences. Also, the performance of treatment sequences AB and BB is currently indistinguishable.

***Example 2*:** Let us consider another two-treatment two-period repeated measurement design for a study of the treatment effect of reducing body temperature when a patient has a fever. Normal body temperature is considered to be 37° C, and a temperature above 37° C is described as a fever. In this example, another set of 10 patients was observed, as shown in Table 1.

In this case, the evaluation function can be defined as a sample deviation from the value of 37° C. A smaller value of g_k indicates a treatment sequence that successfully reduces the body temperature of a patient to normal

$$g_k = \frac{\sum_{i=1}^{2} \sum_j (y_{ijk} - 37)^2}{\|K\|}$$

where y_{ijk} is the body temperature for the j^{th} subject given treatment sequence k in the i^{th} period, and $\|K\|$ is the total number of patients given the treatment sequence k.

In this example, we have $g_{AA} = 0.63$; $g_{AB} = 0.875$; $g_{BA} = 1.75$; and $g_{BB} = 6.25$. Of the four possible treatment sequences, treatment sequence AA performs the best.

Note that the function for *Example 1* basically calculates the mean response, while that for *Example 2* calculates the sample variance for each sequence. However, any function that satisfies the two properties described earlier can be used.

3.2 Selection Criteria

Now we define the optimal selection criteria for our allocation rule.

Consider

$$\Lambda = \lambda \frac{\Theta(\hat{\boldsymbol{A}}_l^k(H_{l-1}))}{\Theta(\hat{\boldsymbol{A}}_l^{k^{(O)}}(H_{l-1}))} + (1-\lambda)\frac{g_{l-1,k}}{g_{l-1,k^{(E)}}}, \quad (2)$$

where $\Theta(\cdot)$ is an optimality criteria function, such as the determinant (D-optimality), the trace (A-optimality) or the maximum eigenvalue (E-optimality) of the information matrix, and $g_{l-1,k}$ is a suitably defined evaluation function based on the first $(l-1)$ patients given treatment sequence k, with $k^{(O)}$, the treatment sequence that maximizes the optimality criteria function $\Theta(\cdot)$, and $k^{(E)}$, the best treatment sequence based on the observed data on the first $(l-1)$ patients under the evaluation function $g_{l-1,k}(\cdot)$. Without loss of generality, we assume that a higher value of $g_{l-1,k}$ indicates a better treatment sequence.

The observed information matrix is $\hat{\boldsymbol{A}}_{l-1}(H_{l-1})$, which is estimated on the basis of data available from the first $(l-1)$ patients using their allocation-and-response history H_{l-1}. The information in this matrix includes: 1) the number of patients assigned to each treatment sequence, and 2) the value of the response variable in each time period for each patient.

The estimated expected Fisher information matrix after the l^{th} observation, given the history of the first $(l-1)$ patients and the assumption that the l^{th} patient will be treated by treatment sequence k, is $\hat{\boldsymbol{A}}_l^k(H_{l-1})$. We can use the plug-in method, where the unknown parameters in $\boldsymbol{A}_l^k(H_{l-1})$ are estimated on the basis of observed data from the first $(l-1)$ patients.

The overall selection criterion Λ has two components. The first component deals with choosing the treatment sequence to maximize the information matrix. The second deals with choosing the treatment sequence that gives the best performance based on the observed data. Prior to the experiment, investigators can choose a parameter $\lambda \in [0,1]$ to balance the two objectives. The value of λ gives weights to these two elements. When $\lambda = 1$, we have the traditional criterion of a statistical optimal design problem. When $\lambda = 0$, all weights are given to choose the design based on the efficacy of treatment. The choice is often driven by what researchers want to emphasize. We will discuss the effect of using different λ values along with specific applications in Section 5. The ultimate goal is to assign a patient to a sequence *k* that maximizes Λ.

To illustrate the utility of the evaluation function and the selection criteria, we consider a matrix representation of model (1)

$$E[\mathbf{y}_{jk}] = \mathbf{X}_k \boldsymbol{\beta}, \quad (3)$$

where $\mathbf{y}_{jk} = (y_{ijk})^T$ is a $p \times 1$ vector of observations from subject j in treatment sequence k, $\boldsymbol{\beta}$ is the column vector of unknown parameters, and $\mathbf{X}_k$ is the design matrix for treatment sequence k.

Then we define the observed information matrix given the data from the first $(l-1)$ patients

$$\hat{\boldsymbol{A}}_{l-1} = \sum_{k \in H_{l-1}} N_k \mathbf{X}_k^T \hat{\mathbf{C}}^{-1} \mathbf{X}_k, \quad (4)$$

where $\hat{\mathbf{C}}$ is the estimated variance-covariance matrix for the response vector $\mathbf{y}_{jk}$.

Under the equicorrelated covariance assumption, one can estimate the variance-covariance matrix using

$$\hat{\mathbf{C}} = \hat{\sigma}_\varepsilon^2 \mathbf{I}_{[p]} + \hat{\sigma}_\xi^2 \mathbf{1}_{[p]} \mathbf{1}_{[p]}^T, \quad (5)$$

where $\hat{\sigma}_\varepsilon^2$ and $\hat{\sigma}_\xi^2$ are estimated using allocation and response history up to the first $(l-1)$ patients, $\mathbf{I}_{[p]}$ is a $p \times p$ identity matrix, and $\mathbf{1}_{[p]}$ is a $p \times 1$ vector of ones.

The estimated information matrix, given the history H_{l-1} and the assumption that the l^{th} patient receiving the treatment sequence k, will become

$$\hat{\boldsymbol{A}}_l^k(H_{l-1}) = \hat{\boldsymbol{A}}_{l-1} + \mathbf{X}_k^T \hat{\mathbf{C}}^{-1} \mathbf{X}_k. \quad (6)$$

3.3 Special Cases

For two-period designs, the design matrix $\mathbf{X}_k$ for a given treatment sequence k becomes

$$\mathbf{X}_{AA} = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 \end{pmatrix},$$

$$\mathbf{X}_{AB} = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & -1 & 1 & 0 \end{pmatrix},$$

$$\mathbf{X}_{BA} = \begin{pmatrix} 1 & 0 & -1 & 0 & 0 \\ 1 & 1 & 1 & -1 & 0 \end{pmatrix},$$

$$\mathbf{X}_{BB} = \begin{pmatrix} 1 & 0 & -1 & 0 & 0 \\ 1 & 1 & -1 & 0 & -1 \end{pmatrix},$$

with the variance-covariance matrix $\mathbf{C}$ of the vector $\mathbf{y}_{jk}$ as in Equation (5) for p=2.

Based on the current observations, H_{l-1}, the estimated information matrix up to the $(l-1)^{th}$ patient as in Equation (4), is

$$\hat{\boldsymbol{A}}_{l-1} = \sum_{k\in H_{l-1}} N_k \mathbf{X}_k^T \hat{\mathbf{C}}^{-1} \mathbf{X}_k$$
$$= \sum_{k\in H_{l-1}} N_k \mathbf{X}_k^T (\hat{\sigma}_\varepsilon^2 \mathbf{I}_{[2]} + \hat{\sigma}_\xi^2 \mathbf{1}_{[2]} \mathbf{1}_{[2]}^T)^{-1} \mathbf{X}_k ,$$

where $\hat{\sigma}_\varepsilon^2$ and $\hat{\sigma}_\xi^2$ are estimated using the restricted maximum likelihood method (Laird and Ware, 1982). Then the estimated information matrix, given the history H_{l-1} and the assumption that the l^{th} patient receiving the treatment sequence k is obtained using Equation (6).

Similarly, for three-period designs, the design matrix $\mathbf{X}_k$ for a given treatment sequence k will be defined as follows.

$$\mathbf{X}_{AAA} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 \end{pmatrix},$$

$$\mathbf{X}_{AAB} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & -1 & 1 & 0 \end{pmatrix},$$

$$\mathbf{X}_{ABA} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & -1 & 1 & 0 \\ 1 & 0 & 1 & 1 & -1 & 0 \end{pmatrix},$$

$$\mathbf{X}_{ABB} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & -1 & 1 & 0 \\ 1 & 0 & 1 & -1 & 0 & -1 \end{pmatrix},$$

$$\mathbf{X}_{BBB} = \begin{pmatrix} 1 & 0 & 0 & -1 & 0 & 0 \\ 1 & 1 & 0 & -1 & 0 & -1 \\ 1 & 0 & 1 & -1 & 0 & -1 \end{pmatrix},$$

$$\mathbf{X}_{BBA} = \begin{pmatrix} 1 & 0 & 0 & -1 & 0 & 0 \\ 1 & 1 & 0 & -1 & 0 & -1 \\ 1 & 0 & 1 & 1 & -1 & 0 \end{pmatrix},$$

$$\mathbf{X}_{BAB} = \begin{pmatrix} 1 & 0 & 0 & -1 & 0 & 0 \\ 1 & 1 & 0 & 1 & -1 & 0 \\ 1 & 0 & 1 & -1 & 1 & 0 \end{pmatrix},$$

$$\mathbf{X}_{BAA} = \begin{pmatrix} 1 & 0 & 0 & -1 & 0 & 0 \\ 1 & 1 & 0 & 1 & -1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \end{pmatrix},$$

along with the covariance matrix and information matrix as obtained from (4)-(6).

4. The Allocation Rule

A new allocation rule for setting up a response adaptive design with a total of N subjects can be conducted as follows for a given λ.

Step 1: The first $m(m < N)$ patients are assigned using the optimal design suggested in the literature or a completely randomized design.

Step 2: To allocate the l^{th} patient, $m+1 \le l \le N$, calculate the observed information matrix based on the data available from the first $(l-1)$ patients, $\hat{A}_{l-1}(H_{l-1})$, and the evaluation function $g_{l-1,k}(H_{l-1})$, where $k = 1, 2, ..., s$.

Step 3: Choose the treatment sequence k^* for the l^{th} patient to maximize the criterion Λ defined in (2). In situations where more than one treatment sequences achieve the maximum criterion score, one can randomly assign one treatment sequence to the l^{th} patient.

Step 4: Repeat steps 2 to 3 until all *N* patients have been allocated.

Note that the above adaptive approach is applicable to both discrete and continuous responses, under suitable model assumptions. In this paper, we implement the allocation rule to construct response adaptive repeated measurement designs with continuous responses.

5. Response Adaptive Repeated Measurement Design

Based on model (1), we construct response adaptive, two-treatment repeated measurement designs. We demonstrate that the efficiency of the designs constructed under the new proposed allocation rule increases with sample size, and these adaptive designs are more efficient than fixed optimal designs in terms of the mean squared error. We also discuss the challenges and partial solutions in generalizing the strategy to multi-treatment and multi-period repeated measurement designs.

We assume that $\sigma_\xi^2 = 2$, $\sigma_\varepsilon^2 = 1$, and $\mu = 100$. We consider the designs for λ =1, 0.9, 0.7, 0.3 and 0 for N =10, 20, 40 and 100, respectively. The $\lambda = 1$ indicates that the only objective is to increase estimation precision, i.e., maximize the information matrix, while $\lambda = 0$ indicates that the only objective is to increase the proportion of patients assigned to a better treatment. When $0 < \lambda < 1$, both objectives are taken into consideration. The adaptive design provides a balanced approach to achieving these two objectives.

Since one of the components of the selection criteria depends on the value of the response, the design also depends on it. This differs from the usual construction of fixed optimal designs.

For $\pi = \tau = \gamma = \varphi = 0$, there is no treatment difference. Table 2 shows the expected outcome, the

mean vector $E[\mathbf{y}_{jk}]$, for each treatment sequence based on the values used for $\pi = \tau = \varphi = 25$ and $\gamma = -25$. Since $\tau = \frac{\tau_A - \tau_B}{2} = 25$, *A* is expected to be better than *B* by 50 in the scale of measurement. Further, $\varphi = \frac{\varphi_A - \varphi_B}{2} = 25$ indicates the self carryover effect of *AA* is expected to produce a better outcome than the self carryover effect of *BB* by 50 units; $\gamma = \frac{\gamma_A - \gamma_B}{2} = -25$ means the mixed carryover effect *AB* is expected to reduce the value of the outcome by 50 units as compared to the mixed carryover effect *BA*; and $\pi = 25$ indicates that the second-period effect is better than the first-period effect, by 25 units.

As mentioned in Section 3, for simplicity we assume that a response of higher magnitude indicates a better treatment and that all responses are nonnegative. Then the summation of all outcomes from a given treatment sequence is a straightforward way of defining an evaluation function. This is the approach taken throughout this section. To smooth out the randomness, we report the average allocation results to treatment sequences from 1,000 repetitions.

5.1 Two-Treatment Two-Period Designs

In two-treatment two-period repeated measurement design, there are four different treatment sequences: *AA*, *AB*, *BA* and *BB*. Suppose the first l patients were assigned using an optimal design suggested in the literature. For example, we can allocate the first l patients equally to the four sequences. We then allocate the rest of the patients adaptively, starting from the observed data from these four patients. Let N_{1l}, N_{2l}, N_{3l} and N_{4l} be the number of patients who have received treatment sequence *AA*, *AB*, *BA* and *BB*, respectively, up to patient l.

For $\pi = \tau = \gamma = \varphi = 0$, there is no treatment difference. Therefore, we expect the design will give equal allocation to all treatment sequences. Table 3 shows, for all combinations of N and λ values, the rule assigns an approximately equal number of subjects to each of the four treatment sequences, as expected. In addition, the estimation of each parameter with its standard error indicates that as sample sizes increase, standard errors in the estimated parameters of interest decrease, as expected (not shown).

Under the treatment difference as specified in Table 2, Table 3 also shows that when $\lambda = 1$, the rule assigns an equal number of subjects to each of the four treatment sequences. This is similar to the traditional fixed optimal design, where the treatment effectiveness is not taken into consideration. Estimation precision is the only consideration in $\lambda = 1$. When $\lambda < 1$, more patients are assigned to treatment sequence *AA*, which is the best combination setting. The rest of the patients, in decreasing order, receive treatments *BA*, *AB* or *BB*. The result is as expected. Based on the values of the parameters of interest, treatment *A* is more effective than treatment *B* ($\tau > 0$); the treatment effect in the second period is stronger than that in the first period ($\pi > 0$); the self carryover effect for treatment sequence *AA* is stronger than for treatment sequence *BB* ($\varphi > 0$), and the mixed carryover effect for treatment sequence *BA* is stronger than that for treatment sequence *AB* ($\gamma < 0$). Therefore, the allocation results are consistent with what one would expect (Table 2).

The estimation of each parameter with its standard error (not shown) indicates that in all cases the estimated values are very close to the true values of the parameters of interest. For a fixed value of λ, the precision of the estimation decreases as N increases. For a fixed value of N, the standard error slightly increases as the value of λ decreases. This happens because when λ decreases we give more emphasis to ethical criteria than to the precision of the estimators. This approach involves a trade-off between benefit and cost.

Afsarinejad and Hedayat (2002) showed that the treatment contrast effect cannot be estimated unbiasedly for any fixed design under the model that allows for two different types of carryover effects from each treatment. Therefore, adaptive designs constructed under the new proposed allocation rule clearly improve the design efficacy for the two-period trials.

5.2. Two-Treatment Three-Period Designs

Eight different treatment sequences are available for assignment. We assume that, at the initial stage, the first l patients are entered into the study. Let N_{kl} be the number of patients receiving treatment sequence k, where $k = AAA$, AAB, ABA, ABB, BBB, BBA, BAB and BAA. Assume that eight subjects were already entered in the study, one for each type of treatment sequence.

Under the traditional model with an equi-correlated covariance structure, the design *ABB* and its dual is known to be the universally optimal design (Laska, Meisner and Kushner 1983, Kershner 1986), while, under the self and mixed carryover effects model, the design ABA and its dual is optimal for estimating the treatment contrast (Hedayat and Stufken, 2003). In this section, we compare the adaptive designs constructed under the new allocation rule to these fixed designs, Design *ABB/BAA* and Design *ABA/BAB*.

As in the two-period two-treatment simulations, we first consider a situation where there are no effects at all. For $\pi_2 = \pi_3 = \tau = \gamma = \varphi = 0$, Table 4 shows that when

$\lambda < 1$, we assign an approximately equal number of subjects to each of the eight treatment sequences, while when $\lambda = 1$, we assign an approximately equal number of subjects to a pair of a treatment sequence and its dual. More subjects were given to *ABB* and its dual, which the fixed optimal design uses for comparing two treatments under the traditional model. However, adaptive designs use six of the eight sequences rather uniformly.

Similar to the two-period two-treatment case, when treatment effects present as $\pi_2 = \pi_3 = \tau = \varphi = 25$ and $\gamma = -25$, the mean vector of each treatment sequence, in decreasing order, is *AAA*, *BAA*, *AAB/ABA*, *BAB/BBA*, *ABB*, and *BBB* (Table 2).

Table 4 shows that, when $\lambda = 1$, we assign an approximately equal number of subjects to a treatment sequence and its dual treatment sequence, and more subjects are given *ABB/BAA*, as before. However, as $\lambda < 1$ and as it decreases, we assign more subjects to a better treatment sequence *AAA* and fewer subjects to the worse treatment sequence *BBB*. This is consistent with what one would expect (Table 2).

Figure 1 shows the estimated relative efficiency (RE) of the adaptive designs relative to Design *ABB/BAA* for estimating $\boldsymbol{\theta} = (\tau, \gamma, \varphi)^T$ under A-, D-, and E-optimality, and for estimating the treatment contrast τ, respectively, where RE > 1 indicates a more efficient design than the reference design. It illustrates that response adaptive designs are more efficient than the fixed optimal design for all λ and optimality criteria. In all cases, adaptive designs increase the design efficiency 1.5 to 4 times as compared to Design *ABB/BAA*. While Design *ABA* and its dual has the highest efficiency for estimating the direct treatment contrast, this design cannot estimate self carryover effects. However, the new proposed adaptive designs are superior to these fixed designs because they can have similar efficiency while also taking the treatment performance into consideration. In all cases, adaptive designs increase design efficiency, especially when the total number of subjects is large. The adaptive designs constructed under the new adaptive allocation rule not only take the treatment performance into consideration, but also have relatively high efficiency.

5.3 Generalization

One can generalize the allocation rule to construct adaptive t-Treatment p-Period repeated measurement designs. The main challenge is to narrow down the number of treatment sequences out of t^p possibilities, which increase substantially as the number of treatments and periods increase. One can consider a particular subset of RMDs, for example, uniform cross-over designs (Bate and Jones, 2002). Alternately, one can refer to the fixed optimal design results available in the literature (Ebbutt 1984, Kershner 1986, Matthews 1987, Carriere and Reinsel 1992&1993, Carriere 1994, Hedayat and Stufken 2003).

For example, Liang (2006) showed that the design $ABBA$ and its dual or $AABA$ and its dual are the optimal two-treatment four-period designs for estimating the treatment contrast based on the self and mixed carryover effects model with random subject effects. Liang (2006) also found that the four-sequence design $ABBA$, $AABB$ and their duals, or the six-sequence design, $ABBA$, $ABAB$, $AABB$ and their duals are the optimal designs for estimating the treatment contrast based on the traditional model. To construct an adaptive design, one can apply the allocation rule to these smaller subsets of all possible treatment sequences.

6. Conclusion

In this paper, we extended single-objective designs to multiple-objective designs by developing a new adaptive allocation rule that can provide efficient estimates of the treatment contrasts and assign more patients to a better treatment. The basic idea is to modify the allocation rule on the basis of observed data from previous patients. We assume patients enter the study sequentially. The first few patients can be assigned using the optimal design suggested in the literature or a completely randomized design. Then the information is updated based on the observed data.

We introduce an evaluation function to evaluate the performance of each treatment sequence. For the incoming set of patients, we consider all possible treatment sequences and choose one that optimizes the criteria, which has two components: 1) to maximize the information matrix and 2) to give the best treatment benefit/performance based on the observed data. It can be easily extended to accommodate any number of other objectives.

Prior to the experiment, the investigator can choose the parameter λ to balance the two objectives. A large value of λ will place more emphasis on estimation precision. When $\lambda = 1$, the allocation rule used in the traditional model will result in response adaptive symmetric designs as considered by Kushner (2003). A small value of λ will emphasize the performance/ benefit of the treatment. When $\lambda = 0$, the allocation rule becomes a typical play-the-winner rule (Zelen, 1969). Note that Kushner's adaptive allocation rule is for trials with continuous outcomes, and Zelen's play-the-winner rule is for trials with dichotomous outcomes. However, our new adaptive allocation rule is applicable to either the trial with continuous or dichotomous outcome.

We utilize this allocation rule to construct adaptive repeated measurement designs with continuous responses/outcome, using the self and mixed carryover effects model. We provide a detailed allocation rule for constructing adaptive two-treatment two- and three-

period repeated measurement designs.

The value of λ is predetermined by researchers and used to balance the two objectives of increasing estimation precision and decreasing the proportion of patients receiving inferior treatments. We find that the design efficiency is a skewed function of λ that decreases sharply as λ decreases. Therefore, choosing a high value of λ is recommended in practice. This results in designs with effective treatment sequences without much loss of estimation precision.

References

1. Afsarinejad, K., and Hedayat, A.S. (2002), "Repeated Measurements Designs for a model with Self and Mixed Carryover Effects," *Journal of Statistical Planning and Inference*, **106**, 449-459.
2. Armitage, P. (1975), *Sequential Medical Trials*, Blackwell, Oxford.
3. Bate, S.T., and Jones, B. (2002), "The Construction of Universally Optimal Uniform Cross-over Designs," *GlaxoSmithKline Pharmaceuticals Technical Report*.
4. Carrière, K.C. (1994a), "Incomplete Repeated Measures Data Analysis in the Presence of Treatment Effects," *Journal of the American Statistical Association*, **89**, 680-686.
5. Carrière, K.C. (1994b), "Cross-over Designs for Clinical Trials," *Statistics in Medicine*, **13**, 1063-1069.
6. Carrière, K.C., and Reinsel, G.C. (1992), "Investigation of Dual-balanced Crossover Designs for Two Treatments," *Biometrics*, **48**, 1157-1164.
7. Carrière, K.C., and Reinsel, G.C. (1993), "Optimal Two-period Repeated Measurement Designs with Two or More Treatments," *Biometrics*, **80**, 924-929.
8. Ebbutt, A.F. (1984), "Three-period Crossover Designs for Two Treatments," *Biometrics*, **40**, 219-224.
9. Hedayat, A., and Afsarinejad, K. (1978), "Repeated Measurements Designs, II," *The Annals of Statistics*, **6(3)**, 619-628.
10. Hedayat, A.S., and Stufken, J. (2003), "Optimal and Efficient Crossover Designs under Different Assumptions about the Carryover Effects", *Journal of Biopharmaceutical Statistics*, **13(3)**, 519-528.
11. Kershner, R.P. (1986), "Optimal 3-period 2-treatment Crossover Designs with and without Baseline Measurements," Proceedings of the Biopharmacerutical Section of the American Statistical Association. 152-156.
12. Kushner, H.B. (2003), "Allocation Rules for Adaptive Repeated Measurements Designs," *Journal of Statistical Planning and Inference*, **113**, 293-313.
13. Laska, E.M., Meisner, M., and Kushner, H.B. (1983), "Optimal Crossover Designs in the Presence of Carryover Effects," *Biometrics*, **39**, 1087-1091.
14. Liang, Y. (2006), *Response-Adaptive Repeated Measurement Design for Clinical Trials*, University of Alberta, PhD thesis.
15. Matthews, J.N.S. (1987), "Optimal Crossover Designs for the Comparison of Two Treatments in the Presence of Carryover Effects and Auto-correlated Errors," *Biometrika*, **74**, 311–320.
16. Moerbeek, M., and Wong, W.K. (2002), "Multiple-objective Optimal Designs for the Hierarchical Linear Model," *Journal of Official Statistics*, **18(2)**, 291–303.
17. Zelen, M. (1969), "Play the Winner Rule and the Controlled Clinical Trial," *Journal of the American Statistical Association*, **64**, 131–146.

Table 1: Example of Data

Subject ID	Treatment Sequence	Period 1		Period 2	
1	*AA*	1	37.5	1	37
2	*AB*	0	37.5	0	38
3	*AA*	1	37.8	1	37
4	*BA*	0	37.5	1	37
5	*BB*	1	38.5	0	39
6	*BA*	0	38.5	0	37.5
7	*AB*	1	37.5	1	37.5
8	*AA*	0	38	1	37
9	*BA*	0	38.5	1	37.5
10	*BB*	1	39	0	38.5

Note: Entries in the shaded columns are data for a dichotomous response and those in the columns to the right are data for a continuous response.

Table 2: Expected Outcome for Each Treatment Sequence Based on the Values Used for Simulations

Design	Value of Parameter	Treatment Sequence	Expected Outcomes
Two-Treatment Two-Period	$\mu = 100$ $\pi = \tau = \varphi = 25$ $\gamma = -25$	*AA* *AB* *BA* *BB*	$(125, 175)^T$ $(125, 75)^T$ $(75, 175)^T$ $(75, 75)^T$
Two-Treatment Three-Period	$\mu = 100$ $\pi = \tau = \varphi = 25$ $\gamma = -25$	*AAA* *AAB* *ABA* *ABB* *BBB* *BBA* *BAB* *BAA*	$(125, 175, 175)^T$ $(125, 175, 75)^T$ $(125, 75, 175)^T$ $(125, 75, 75)^T$ $(75, 75, 75)^T$ $(75, 75, 175)^T$ $(75, 175, 75)^T$ $(75, 175, 175)^T$

Note: Entries are the expected mean vectors $\mathbf{y}_{jk}$.

Table 3: Estimated Numbers of Patients for Each Treatment Sequence Using a Two-Period Design

N	λ	N_{AA}		N_{AB}		N_{BA}		N_{BB}	
10	1	2.502	2.519	2.498	2.481	2.466	2.441	2.534	2.559
	0.9	2.497	3.000	2.503	2.000	2.510	3.000	2.490	2.000
	0.7	2.501	3.000	2.499	2.000	2.488	3.000	2.512	2.000
	0.3	2.479	3.000	2.518	2.000	2.491	3.000	2.512	2.000
	0	2.527	3.000	2.506	2.000	2.484	3.000	2.483	2.000
40	1	10.000	10.000	10.000	10.000	10.000	10.000	10.000	10.000
	0.9	10.000	12.000	10.000	9.000	10.000	11.000	10.000	8.000
	0.7	10.000	13.000	10.000	9.000	10.000	11.001	10.000	6.999
	0.3	10.000	13.416	10.000	8.986	10.000	11.030	10.000	6.568
	0	10.000	13.717	10.000	8.979	10.000	11.022	10.000	6.282
80	1	20.000	20.000	20.000	20.000	20.000	20.000	20.000	20.000
	0.9	19.998	25.547	20.007	17.930	20.000	22.405	19.995	14.118
	0.7	19.997	26.810	20.009	17.693	19.999	22.451	19.995	13.046
	0.3	20.001	27.127	19.995	17.624	20.002	22.426	20.002	12.823
	0	19.995	27.243	20.000	17.602	19.998	22.374	20.007	12.781
100	1	25.000	25.000	25.000	25.000	25.000	25.000	25.000	25.000
	0.9	25.008	32.299	24.991	22.197	25.001	28.103	25.000	17.401
	0.7	24.999	33.664	24.999	22.021	24.997	28.089	25.005	16.226
	0.3	25.004	34.051	24.986	21.959	24.997	28.054	25.013	15.936
	0	24.971	34.134	25.014	21.955	25.006	28.051	25.009	15.860

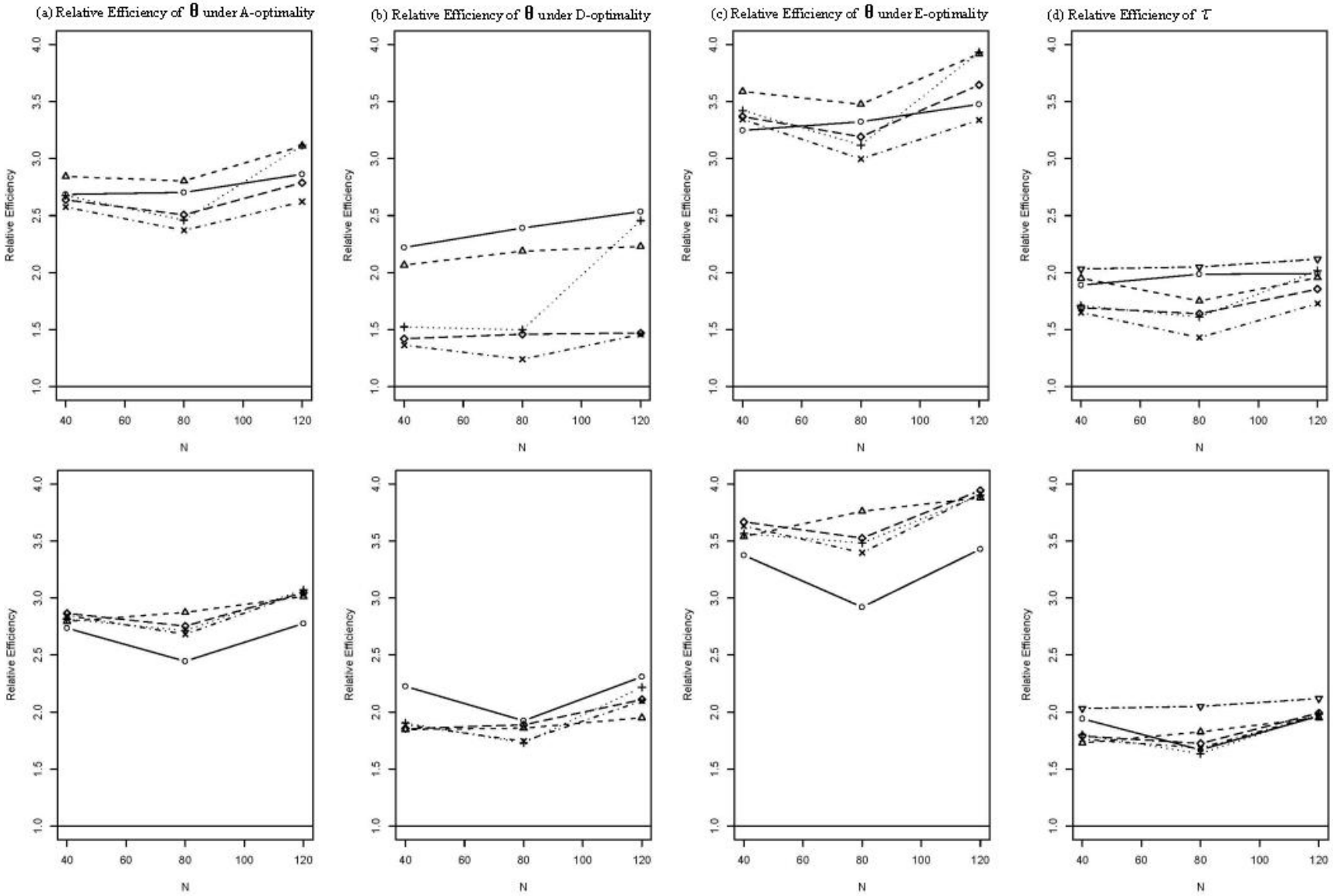
Note: Entries in the shaded columns are based on 1,000 computer replications under 2-treatment 2-period RMDs with $\pi = \tau = \gamma = \varphi = 0$, and those to the right are for $\pi = \tau = \varphi = 25$, $\gamma = -25$, for $\sigma_{\xi}^2 = 2$, $\sigma_{\varepsilon}^2 = 1$, and $\mu = 100$.

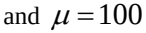
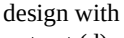
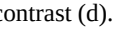
Table 4: Estimated Numbers of Patients for Each Treatment Sequence Using a Three-Period Design

N	λ	N_{AAA}	N_{AAB}	N_{ABA}	N_{ABB}	N_{BBB}	N_{BBA}	N_{BAB}	N_{BAA}
40	1	1.016	5.984	5.977	7.023	1.018	5.982	5.974	7.026
		1.014	5.986	5.968	7.032	1.014	5.986	5.980	7.020
	0.9	4.144	5.000	5.591	5.263	4.149	5.000	5.609	5.244
		5.144	5.000	6.000	4.856	3.000	5.000	5.000	6.000
	0.7	5.000	5.000	5.000	5.000	5.000	5.000	5.000	5.000
		6.009	5.003	5.977	4.000	3.011	5.000	5.000	6.000
	0.3	5.000	5.000	5.000	5.000	5.000	5.000	5.000	5.000
		6.975	5.001	5.010	4.000	3.014	5.000	5.000	6.000
	0	5.000	5.000	5.000	5.000	5.000	5.000	5.000	5.000
		6.996	5.000	5.000	4.000	3.005	5.000	4.999	6.000
80	1	1.009	12.998	11.848	14.145	1.007	13.000	11.842	14.151
		1.006	12.998	11.842	14.154	1.009	12.995	11.839	14.157
	0.9	9.065	9.996	10.647	10.297	9.062	10.000	10.628	10.305
		12.001	10.28	11.038	8.736	6.328	9.501	9.999	12.117
	0.7	10.000	10.000	10.001	10.000	9.998	10.000	10.001	10.000
		12.999	10.81	10.996	8.001	6.749	9.119	9.321	12.005
	0.3	10.000	10.000	10.000	10.000	10.000	10.000	10.000	10.000
		13.067	10.947	10.973	8.000	6.859	9.071	9.083	12.000
	0	10.000	10.000	10.000	10.000	10.000	10.000	10.000	10.000
		13.151	10.944	10.947	8.000	6.840	9.057	9.060	12.001
120	1	1.010	20.016	17.649	21.325	1.007	20.019	17.630	21.344
		1.011	20.021	17.645	21.323	1.008	20.024	17.651	21.317
	0.9	14.098	14.972	15.629	15.324	14.089	14.964	15.592	15.332
		18.662	15.878	16.712	12.310	9.767	14.008	14.444	18.219
	0.7	14.950	14.994	15.039	15.019	14.953	14.999	15.028	15.018
		19.968	15.999	16.029	11.998	9.971	13.999	14.007	18.029
	0.3	14.997	14.999	15.000	15.004	14.999	14.996	15.002	15.003
		20.010	15.998	16.005	11.999	9.98	14.000	14.000	18.008
	0	15.001	14.996	15.003	15.003	14.999	15.004	14.998	14.996
		20.019	16.000	16.000	11.992	9.988	13.999	13.998	18.004

Note: Entries are based on 1,000 computer replications under 2-treatment 3-period RMDs with $\pi_2 = \pi_3 = \tau = \gamma = \varphi = 0$ (shaded) and $\pi_2 = \pi_3 = \tau = \varphi = 25$, $\gamma = -25$, respectively with $\sigma_\xi^2 = 2$, $\sigma_\varepsilon^2 = 1$, and $\mu = 100$.

Figure 1: Relative Efficiency of the Parameters of Interest Compared to Design *ABB/BAA* for Two-treatment Three-period RMDs



Note: Relative efficiencies in the upper row are calculated by using 1,000 computer replications with $\pi_2 = \pi_3 = \tau = \gamma = \varphi = 0$, the bottom row with $\pi = \tau = \varphi = 25$, $\gamma = -25$, for $\sigma_{\xi}^2 = 2$, $\sigma_{\varepsilon}^2 = 1$, and $\mu = 100$, where  adaptive design with λ =1; adaptive design with λ =0.9;  adaptive design with λ =0.7;  adaptive design with λ =0.3; adaptive design with λ =0; and design *ABA/BAB*. Since Design *ABA/BAB* cannot estimate self carryover effect contrast, the efficiency is compared only for the direct treatment effect contrast (d).