Published in final edited form as:

J Stat Plan Inference. 2017 March ; 182: 88–99. doi:10.1016/j.jspi.2016.09.002.

Implementing Optimal Allocation in Clinical Trials with Multiple Endpoints

Lu Wang^a, Yong Chen^b, and Hongjian Zhu^{a,*}

^aDepartment of Biostatistics, University of Texas Health Science Center School of Public Health at Houston, 1200 Pressler St, Houston, Texas 77030, USA

^bDepartment of Biostatistics and Epidemiology, University of Pennsylvania, 210 Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104

Abstract

Modern clinical trials are often complex, with multiple competing objectives and multiple endpoints. Such trials should be both ethical and efficient. In this paper, we overcome the obstacles introduced by the large number of unknown parameters and the possible correlations between the multiple endpoints. We obtain the optimal allocation proportions for the following two optimization problems: (1) maximizing the power of the test of homogeneity with a fixed sample size, and (2) minimizing the expected weighted number of failures with a fixed power. Further, we implement these optimal allocations through response-adaptive randomization procedures. Our theoretical results provide the foundation for the implementation and further investigation of the procedure, and our numerical studies demonstrate its ability to achieve diverse objectives.

Keywords

Clinical trial; Multiple endpoints; Optimal allocation; Power; Response-adaptive design

1. Introduction

Clinical trials are complex experiments involving human beings. They often have multiple competing objectives, such as minimizing the total number of failures and maximizing the power of detecting treatment effects. These cannot be satisfied by traditional clinical trial designs that aim to balance the patient numbers across different treatments. Response-adaptive randomization (RAR) procedures achieve these objectives by skewing the allocation probability according to the previous treatment assignments and responses. Most of the research on RAR designs has been focused on clinical trials with a single endpoint. However, in modern clinical trials multiple possibly correlated endpoints are often evaluated

*Corresponding author. Tel: 1(713)500-9580. hongjian.zhu@uth.tmc.edu (Hongjian Zhu).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

simultaneously, and the clinical goal is to show a positive effect for at least one endpoint. There are different rationales for multiple (composite) endpoints. First, several different endpoints may be important for the participants. Second, the investigators are not always sure which outcome indicates a treatment effect (Friedman *et al.*, 2010). Third, embedding multiple endpoints into the analysis can help to distinguish weaker signals of treatment effects from the noises of sampling errors (Moyé, 2003). Fourth, the total event rate can be increased, which can lead to a reduction in the required sample size.

Numerous real trials currently use multiple endpoints. For example, pulmonary function, neuro-psychological status, quality of life, and mortality were all assessed in the Nocturnal Oxygen Therapy Trial (1980); the trial compared continuous oxygen (O₂) therapy and 12-hour nocturnal O₂ therapy for chronic obstructive pulmonary disease. The Urokinase Pulmonary Embolism Trial (1974) explored the effect of three treatments, 12 hours of urokinase, 24 hours of urokinase, and 24 hours of streptokinase, on pulmonary embolism. They considered three endpoints: angiographic severity, lung scan perfusion defects, and hemodynamic variables. Roberts *et al.* (1984) investigated the effect of propranolol in limiting the myocardial infarct size; their endpoints were the precordial electrocardiogram mapping, radionuclide studies, and enzyme levels. In a clinical trial of an analgesic drug used to relieve arthritic pain (Jennison and Turnbull, 1993), the primary endpoints were a measure of the pain relief experienced by the patient and the effect on the arthritic condition of the joint. Since the drug's success in relieving pain may lead the patient to be less careful of the joint, the two outcome measures might be related. These examples motivate our investigation of the use of RAR designs to find optimal allocations for clinical trials with multiple endpoints; the goal is to achieve ethical and efficient objectives.

The idea underlying RAR designs can be traced back to Thompson (1933) and Robbins (1952). The play-the-winner rule (Zelen, 1969) and the randomized play-the-winner rule (Wei and Durham, 1978) can reduce the number of patients receiving inferior treatments. Rosenberger *et al.* (2001) proposed an optimal allocation that minimizes the total number of failures while fixing the power. Ivanova and Rosenberger (2000) showed that an unequal allocation can sometimes result in a gain in the power. Ivanova (2003) proposed the drop-the-loser rule that achieves minimal variability. Hu and Rosenberger (2003) proved that RAR designs can increase statistical efficiency in certain clinical trials. A comprehensive introduction to RAR designs can be found in Hu and Rosenberger (2006).

Hu and Rosenberger (2003) formalized a three-step approach for the development of optimal RAR procedures. The first step is to mathematically formulate the objectives, such as maximizing the power, and to derive an optimal allocation given these objectives. For example, when comparing two binary responses with success rates p_A and p_B , given the objective of maximizing the power, the optimal allocation proportions η_A and η_B for the two treatments will be the Neyman allocation:

$$\eta_A = \frac{\sqrt{p_A(1-p_A)}}{\sqrt{p_A(1-p_A)} + \sqrt{p_B(1-p_B)}} \text{ and } \eta_B = \frac{\sqrt{p_B(1-p_B)}}{\sqrt{p_A(1-p_A)} + \sqrt{p_B(1-p_B)}}.$$

The second step is to develop RAR procedures such as the doubly adaptive biased coin design (DBCD) proposed by Hu and Zhang (2004) to target the theoretically optimal allocation derived in the first step. The third step is to study the operating characteristics of the proposed procedure. It is clear that the derivation of the optimal allocation for certain goals is essential for RAR procedures. In addition, the closed form of the theoretically optimal allocation plays a crucial role in evaluating the properties of the RAR procedure such as the lower bound of the variability (Hu *et al.*, 2006), although the numerical results are sufficient to apply the procedure in practice.

A general framework for the optimization problem was given by Jennison and Turnbull (1999), triggering a series of publications on optimal RAR designs with binary outcomes (Rosenberger *et al.*, 2001; Ivanova and Rosenberger, 2001; Rosenberger and Hu, 2004; Jeon and Hu, 2012), normal outcomes (Biswas and Mandal, 2004; Zhang and Rosenberger, 2006; Gwise *et al.*, 2008; Biswas and Bhattacharya, 2009, 2010), and survival outcomes (Zhang and Rosenberger, 2007). Tyomfyeyev *et al.* (2007) established a general mathematical framework for obtaining optimal allocations using the Karush-Kuhn-Tucker (KKT) conditions (Kuhn and Tucker, 1951; Karush, 1939). However, all these studies explored trials with a single endpoint.

We investigate optimal allocation for clinical trials with two possibly correlated endpoints, and we implement this using RAR designs. We study how to maximize the power of a test of homogeneity for a fixed sample size and how to minimize the expected weighted number of failures (EWNF) for a fixed power. We first obtain the analytical solution of the optimal allocations for these problems. It is not trivial to check the applicability of the KKT conditions in the presence of multiple endpoints with a covariance structure. This leads to several higher-degree equations that are difficult to solve analytically. The numerical solutions are often sufficient to implement the RAR procedure. However, without the derived closed form, it is difficult to study further theoretical properties of the RAR designs, the third step of the process (Hu and Rosenberger, 2003). This is one of the major contributions of this paper. Further, we have performed comprehensive numerical studies based on simulated data and the redesign of a clinical trial. Our method is shown to be more ethical and efficient than traditional designs. The power is moderately increased and the expected number of failures among trial participants is reduced using the RAR designs for the optimal allocations. Therefore, both our theoretical and numerical results provide important insight for implementing and further investigating RAR designs in clinical trials with multiple endpoints.

This article is organized as follows. In Section 2, we present the framework, define the two optimization problems, and derive the optimal allocation proportions. In Section 3, we implement the optimal allocation using the DBCD proposed by Hu and Zhang (2004) and report the results of our numerical studies. We give a discussion in Section 4, and the proof of Theorems is in the Appendix.

2. Optimization problems in clinical trials with two endpoints

2.1. Framework, notation, and optimization problems

Consider a randomized clinical trial with two treatment groups and two possibly correlated binary endpoints. Assume that there are n_A patients in treatment A and n_B patients in treatment B at the end of the trial, and let $\mathbf{n} = (n_A, n_B)$. Let (p_{A1}, p_{A2}) and (p_{B1}, p_{B2}) be the success rates of the two endpoints in treatments A and B, respectively, and ρ_A and ρ_B be the correlations between these two endpoints for treatments A and B, respectively.

We wish to test the hypothesis

$$H_0: \mathbf{p} = \mathbf{0} \quad \text{vs} \quad H_A: \mathbf{p} \neq \mathbf{0}$$

where $\mathbf{p} = (p_{A1} - p_{B1}, p_{A2} - p_{B2})$. Clinical trials are complex and have different types of clinical decision rules, inferential goals, and common approaches to analyze multiple endpoints (Dmitrienko et al., 2009), among which one important rule is that the trial's outcome would be declared positive if at least one endpoint was significantly improved compared to the control. One motivating example is the STRIVE study (Zeiher et al. 2004) which aimed to determine whether sivelestat would reduce 28-day all-cause mortality or increase the number of ventilator-free days compared with placebo in mechanically ventilated patients with acute lung injury. This paper focuses on this scenario, and the methodology that is to follow will be applied to redesign a specific study in Section 3.2. It is worth noting that there are other clinical decision rules. For example, in some trials, a clinically meaningful effect is defined as the simultaneous improvement in multiple endpoints, and multiple one-sided alternative hypotheses will be tested. Bretz et al. (2009) discussed graphical approach when endpoints are tested in a hierarchical manner. These are beyond the scope of the paper.

Let $\hat{\mathbf{p}} = (\hat{p}_{A1} - \hat{p}_{B1}, \hat{p}_{A2} - \hat{p}_{B2})$ be the maximum likelihood estimator (MLE) of $\mathbf{p}$ and $\hat{\Sigma}$ the estimator of $\Sigma = \text{Var}(\hat{\mathbf{p}})$. Here, Σ can be written as

$$\Sigma = \begin{bmatrix} \frac{p_{A1}q_{A1}}{n_A} + \frac{p_{B1}q_{B1}}{n_B} & P \\ P & \frac{p_{A2}q_{A2}}{n_A} + \frac{p_{B2}q_{B2}}{n_B} \end{bmatrix},$$

where $q_{A1} = 1 - p_{A1}$, $q_{A2} = 1 - p_{A2}$, $q_{B1} = 1 - p_{B1}$, $q_{B2} = 1 - p_{B2}$, and P is defined to be $\rho_A \sqrt{p_{A1}q_{A1}p_{A2}q_{A2}}/n_A + \rho_B \sqrt{p_{B1}q_{B1}p_{B2}q_{B2}}/n_B$, the covariance of $(\hat{p}_{A1} - \hat{p}_{B1})$ and $(\hat{p}_{A2} - \hat{p}_{B2})$. We use $\hat{\mathbf{p}}\hat{\Sigma}^{-1}\hat{\mathbf{p}}'$ as the test statistic; it converges to χ^2_2 under the null hypothesis. Under the alternative hypothesis, it follows a noncentral chi-squared distribution $\chi^2_2(\phi(\mathbf{n}))$ with noncentrality parameter $\phi(\mathbf{n})$ defined as

$$\begin{aligned} \phi(\mathbf{n}) &= \mathbf{p}' \Sigma^{-1} \mathbf{p} \\ &= \left\{ (p_{A1} - p_{B1})^2 \left(\frac{p_{A2} q_{A2}}{n_A} + \frac{p_{B2} q_{B2}}{n_B} \right) + (p_{A2} - p_{B2})^2 \left(\frac{p_{A1} q_{A1}}{n_A} + \frac{p_{B1} q_{B1}}{n_B} \right) \right. \\ &\quad \left. - 2P(p_{A1} - p_{B1})(p_{A2} - p_{B2}) \right\} / \left\{ \left(\frac{p_{A1} q_{A1}}{n_A} + \frac{p_{B1} q_{B1}}{n_B} \right) \left(\frac{p_{A2} q_{A2}}{n_A} + \frac{p_{B2} q_{B2}}{n_B} \right) - P^2 \right\}. \quad (1) \end{aligned}$$

Note that the power can be expressed as an increasing function of the non-centrality parameter $\phi(\mathbf{n})$ (Patnaik, 1949).

We now formulate two equivalent optimization problems. The first maximizes the power of the test by maximizing the above noncentrality parameter under certain conditions on the weighted sum of the sample sizes:

$$\begin{aligned} &\max_{n_A, n_B} \phi(n_A, n_B) \\ &\text{such that } n_A, n_B \geq D(n_A + n_B), \omega_A n_A + \omega_B n_B \leq M. \quad (2) \end{aligned}$$

Here M is a positive constant, the weight ω_A and ω_B are also positive, and the constant $D \in [0, 1/2]$ is the lower bound for the allocation proportion $n_A/(n_A + n_B)$ and $n_B/(n_A + n_B)$, that allows us to control the feasible region and avoid having too few people in either of the treatments. Our goal is to find the allocation proportion, $\eta_A^* = n_A^*/(n_A^* + n_B^*)$ and $\eta_B^* = 1 - \eta_A^*$, as the solution of the optimization problem. The second optimization problem minimizes the weighted sum of the sample sizes under certain conditions on the noncentrality parameter:

$$\begin{aligned} &\min_{n_A, n_B} \omega_A n_A + \omega_B n_B \\ &\text{such that } n_A, n_B \geq D(n_A + n_B), \phi(n_A, n_B) \geq C \quad (3) \end{aligned}$$

where C is a positive constant and D is the constant introduced above. Here, the positive M and C are used to define the optimization problem, and our final solutions with regard to the allocation proportions instead of sample size do not depend on M or C . The value of D is less important for clinical trials with two treatments than those with multiple treatments where $D = 0$ may lead to solutions on the boundary that involves only the best and the worst treatments.

The equivalence of the above optimization problems with regard to the same allocation proportions was discussed by Tymofeyev *et al.* (2007). We prove that the noncentrality parameter $\phi(n_A, n_B)$ in (1) is a concave function, and $\nabla \phi \neq 0$ in the Appendix. We can vary the optimization problems by changing the weights $\omega = (\omega_A, \omega_B)'$. For example, setting $\omega = (1, 1)'$ in problem (2) maximizes the power for a fixed sample size; setting $\omega = (q_{A1} + q_{A2}, q_{B1} + q_{B2})'$ in problem (3) minimizes the expected number of failure from two endpoints for a fixed power.

2.2. Maximizing the power for a fixed sample size

We first consider the optimal allocation to maximize the power for a fixed sample size.

Theorem 1—Let $r = \eta_A^* / \eta_B^*$ and $\tilde{D} = D/(1 - D)$. Then the solution for problem (2) with weights $\omega = (\omega_A, \omega_B)' = (1, 1)'$ is as follows:

1. If $p_{A1} = p_{B1}$, $p_{A2} = p_{B2}$, η_A^* can be any number between 0 and 1 greater than D, and r can be any number between $\tilde{D}$ and $\tilde{D}^{-1}$.
2. If $(p_{A1} - p_{B1})^2 + (p_{A2} - p_{B2})^2 = 0$, let

$$H = b^2 - ac, I = ae - 4bd + 3c^2, G = a^2d - 3abc + 2b^3, J = \frac{4H^3 - a^2HI - G^2}{a^3}$$

where

$$a = (p_{A1} - p_{B1})^2 p_{B1} q_{B1} p_{B2}^2 q_{B2}^2 (1 - \rho_B^2) + (p_{A2} - p_{B2})^2 p_{B1}^2 q_{B1}^2 p_{B2} q_{B2} (1 - \rho_B^2) - 2(p_{A1} - p_{B1})(p_{A2} - p_{B2})(p_{B1} q_{B1} p_{B2} q_{B2})^{\frac{3}{2}} \rho_B (1 - \rho_B^2)$$

$$b = \frac{1}{2}(p_{A1} - p_{B1})^2 p_{A2} q_{A2} p_{B1} q_{B1} p_{B2} q_{B2} (1 - \rho_B^2) + \frac{1}{2}(p_{A2} - p_{B2})^2 p_{A1} q_{A1} p_{B1} q_{B1} p_{B2} q_{B2} (1 - \rho_B^2) - (p_{A1} - p_{B1})(p_{A2} - p_{B2}) \sqrt{p_{A1} q_{A1} p_{A2} q_{A2} p_{B1} q_{B1} p_{B2} q_{B2}} \rho_A (1 - \rho_B^2)$$

$$c = \left[(p_{A1} - p_{B1})^2 \{ p_{A2}^2 q_{A2}^2 p_{B1} q_{B1} - p_{A1} q_{A1} p_{B2}^2 q_{B2}^2 + p_{A1} q_{A1} p_{A2} q_{A2} p_{B2} q_{B2} \rho_A^2 - p_{A2} q_{A2} p_{B1} q_{B1} p_{B2} q_{B2} \rho_B^2 - 2(p_{A2} q_{A2} - p_{B2} q_{B2}) \sqrt{p_{A1} q_{A1} p_{A2} q_{A2} p_{B1} q_{B1} p_{B2} q_{B2}} \rho_A \rho_B \} + (p_{A2} - p_{B2})^2 \{ p_{A1}^2 q_{A1}^2 p_{B2} q_{B2} - p_{A2} q_{A2} p_{B1}^2 q_{B1}^2 + p_{A1} q_{A1} p_{A2} q_{A2} p_{B1} q_{B1} \rho_A^2 - p_{A1} q_{A1} p_{B1} q_{B1} p_{B2} q_{B2} \rho_B^2 - 2(p_{A1} q_{A1} - p_{B1} q_{B1}) \sqrt{p_{A1} q_{A1} p_{A2} q_{A2} p_{B1} q_{B1} p_{B2} q_{B2}} \rho_A \rho_B \} - 2(p_{A1} - p_{B1})(p_{A2} - p_{B2}) \times \{ (p_{A1} q_{A1} p_{B2} q_{B2} + p_{A2} q_{A2} p_{B1} q_{B1} + p_{B1} q_{B1} p_{B2} q_{B2}) \sqrt{p_{A1} q_{A1} p_{A2} q_{A2}} \rho_A - (p_{A1} q_{A1} p_{B2} q_{B2} + p_{A2} q_{A2} p_{B1} q_{B1} + p_{A1} q_{A1} p_{A2} q_{A2}) \sqrt{p_{B1} q_{B1} p_{B2} q_{B2}} \rho_B - p_{A1} q_{A1} p_{A2} q_{A2} \sqrt{p_{B1} q_{B1} p_{B2} q_{B2}} \rho_A^2 \rho_B + p_{B1} q_{B1} p_{B2} q_{B2} \sqrt{p_{A1} q_{A1} p_{A2} q_{A2}} \rho_A \rho_B^2 \} \right] / 6$$

$$d = -\frac{1}{2}(p_{A1} - p_{B1})^2 p_{A1} q_{A1} p_{A2} q_{A2} p_{B2} q_{B2} (1 - \rho_A^2) - \frac{1}{2}(p_{A2} - p_{B2})^2 p_{A1} q_{A1} p_{A2} q_{A2} p_{B1} q_{B1} (1 - \rho_A^2) + (p_{A1} - p_{B1})(p_{A2} - p_{B2}) p_{A1} q_{A1} p_{A2} q_{A2} \sqrt{p_{B1} q_{B1} p_{B2} q_{B2}} \rho_B (1 - \rho_A^2)$$

$$e = -(p_{A1} - p_{B1})^2 p_{A1} q_{A1} p_{A2}^2 q_{A2}^2 (1 - \rho_A^2) - (p_{A2} - p_{B2})^2 p_{A1}^2 q_{A1}^2 p_{A2} q_{A2} (1 - \rho_A^2) + 2(p_{A1} - p_{B1})(p_{A2} - p_{B2})(p_{A1} q_{A1} p_{A2} q_{A2})^{\frac{3}{2}} \rho_A (1 - \rho_A^2).$$

$$\text{i. If } G = 0, r = \max \left((-b + \sqrt{3H + \sqrt{12H^2 - a^2I}}) / a, \tilde{D} \right).$$

- ii. If $G = 0, I = J = 0, r = \max \left(\left\{ -b - 3 \operatorname{sign}(G) \sqrt{H} \right\} / a, \tilde{D} \right)$.
- iii. Let $\Delta = \tilde{P}^2 - 27\tilde{J}^2$. If $G = 0, \tilde{P}^2 + \tilde{J}^2 = 0, \Delta < 0$,
 $r = \max \left((-b - \operatorname{sign}(G) \sqrt{t} + \sqrt{|G|} / \sqrt{t} - t + 3H) / a, \tilde{D} \right)$, where
 $t = \frac{a}{2} \left(\sqrt[3]{-J + \sqrt{-\Delta/27}} + \sqrt[3]{-J - \sqrt{-\Delta/27}} + H \right)$
- iv. If $G = 0, \tilde{P}^2 + \tilde{J}^2 = 0, \Delta = 0$,
 $r = \max \left((-b - \operatorname{sign}(G) \sqrt{y_1} + \sqrt{y_2} + \sqrt{y_3}) / a, \tilde{D} \right)$, where

$$y_1 = a \sqrt{\frac{|I|}{3}} \cos \left(\frac{\theta}{3} \right) + H, \quad y_{2,3} = a \sqrt{\frac{|I|}{3}} \cos \left(\frac{\theta}{3} \pm \frac{2\pi}{3} \right) + H,$$

$$\theta = \cos^{-1} \left(\frac{-J}{\sqrt{|I|^3/27}} \right).$$

Remark 1: When there is no difference between the two groups for one endpoint and no correlation between the endpoints, the allocation reduces to the Neyman allocation. That is, when $p_{A1} = p_{B1}, p_{A2} = p_{B2}$, and $\rho_A = \rho_B = 0$, we have
 $r = n_A^* / n_B^* = \max \left(\sqrt{p_{A2} q_{A2}} / \sqrt{p_{B2} q_{B2}}, \tilde{D} \right)$. Because we assign patients to the different treatments based on just one endpoint it is natural that we get the same result as for the one-endpoint problem.

2.3. Minimizing the expected number of treatment failures for a fixed power

Next, we consider the optimal allocation to minimize the EWNF for a fixed power. We assign different weights to the two endpoints to emphasize the importance of the individual endpoint.

Let $w_j = 0, j = 1, 2$, be the weight for endpoint j . Then the EWNF is

$$w_1 q_{A1} n_A + w_1 q_{B1} n_B + w_2 q_{A2} n_A + w_2 q_{B2} n_B = (w_1 q_{A1} + w_2 q_{A2}) n_A + (w_1 q_{B1} + w_2 q_{B2}) n_B.$$

This is problem (3) with $\omega = (w_1 q_{A1} + w_2 q_{A2}, w_1 q_{B1} + w_2 q_{B2})'$. When $w_1 = w_2 = 1$, the two endpoints are equally important.

Theorem 2—Let $r = \eta_A^* / \eta_B^*$ and $\tilde{D} = D / (1 - D)$. Then the solution for problem (3) with $\omega = (w_1 q_{A1} + w_2 q_{A2}, w_1 q_{B1} + w_2 q_{B2})'$ is as follows:

1. If $p_{A1} = p_{B1}, p_{A2} = p_{B2}$, η_A^* can be any number between 0 and 1 greater than D , and r can be any number between $\tilde{D}$ and $\tilde{D}^{-1}$.
2. If $(p_{A1} - p_{B1})^2 + (p_{A2} - p_{B2})^2 = 0$, let $\omega_A = w_1 q_{A1} + w_2 q_{A2}$ and $\omega_B = w_1 q_{B1} + w_2 q_{B2}$. Then the optimal allocations have the same form as in Theorem 1 with

$$a = \omega_A (p_{A1} - p_{B1})^2 p_{B1} q_{B1} p_{B2}^2 q_{B2}^2 (1 - \rho_B^2) \\ + \omega_A (p_{A2} - p_{B2})^2 p_{B1}^2 q_{B1}^2 p_{B2} q_{B2} (1 - \rho_B^2) \\ - 2\omega_A (p_{A1} - p_{B1})(p_{A2} - p_{B2})(p_{B1} q_{B1} p_{B2} q_{B2})^{\frac{3}{2}} \rho_B (1 - \rho_B^2)$$

$$b = \frac{1}{2}\omega_A (p_{A1} - p_{B1})^2 p_{B1} q_{B1} p_{A2} q_{A2} p_{B2} q_{B2} (1 - \rho_B^2) \\ + \frac{1}{2}\omega_A (p_{A2} - p_{B2})^2 p_{B1} q_{B1} p_{A1} q_{A1} p_{B2} q_{B2} (1 - \rho_B^2) \\ - (p_{A1} - p_{B1})(p_{A2} - p_{B2}) p_{B1} q_{B1} p_{B2} q_{B2} \sqrt{p_{A1} q_{A1} p_{A2} q_{A2}} \rho_A \omega_A (1 - \rho_B^2)$$

$$c = \left((p_{A1} - p_{B1})^2 \{ -\omega_B p_{A1} q_{A1} p_{B2}^2 q_{B2}^2 - \omega_A p_{B1} q_{B1} p_{A2}^2 q_{A2}^2 + p_{A1} q_{A1} p_{A2} q_{A2} p_{B2} q_{B2} \rho_A^2 \omega_A \right. \\ \left. - p_{A2} q_{A2} p_{B1} q_{B1} p_{B2} q_{B2} \rho_B^2 \omega_B \right. \\ \left. - 2(p_{A2} q_{A2} \omega_A - p_{B2} q_{B2} \omega_B) \sqrt{p_{A1} q_{A1} p_{A2} q_{A2} p_{B1} q_{B1} p_{B2} q_{B2}} \rho_A \rho_B \right\} \\ + (p_{A2} - p_{B2})^2 \{ -\omega_B p_{A2} q_{A2} p_{B1}^2 q_{B1}^2 + \omega_A p_{B2} q_{B2} p_{A1}^2 q_{A1}^2 + p_{A1} q_{A1} p_{A2} q_{A2} p_{B1} q_{B1} \rho_A^2 \omega_A \\ - p_{A1} q_{A1} p_{B1} q_{B1} p_{B2} q_{B2} \rho_B^2 \omega_B \\ - 2(p_{A1} q_{A1} \omega_A - p_{B1} q_{B1} \omega_B) \sqrt{p_{A1} q_{A1} p_{A2} q_{A2} p_{B1} q_{B1} p_{B2} q_{B2}} \rho_A \rho_B \} \\ - 2[\{ (p_{A2} q_{A2} p_{B1} q_{B1} + p_{A1} q_{A1} p_{B2} q_{B2}) \omega_A \rho_A + p_{B1} q_{B1} p_{B2} q_{B2} \omega_B \rho_B (1 + \rho_A^2) \} \times \sqrt{p_{A1} q_{A1} p_{A2} q_{A2}} \\ - \{ (p_{A2} q_{A2} p_{B1} q_{B1} + p_{A1} q_{A1} p_{B2} q_{B2}) \omega_B \rho_B + p_{A1} q_{A1} p_{A2} q_{A2} \omega_A \rho_A (1 + \rho_B^2) \} \times \sqrt{p_{B1} q_{B1} p_{B2} q_{B2}}] (p_{A1} - p_{B1})(p_{A2} - p_{B2})) / 6$$

$$d = -\frac{1}{2}\omega_B (p_{A1} - p_{B1})^2 p_{A1} q_{A1} p_{A2} q_{A2} p_{B2} q_{B2} (1 - \rho_A^2) \\ - \frac{1}{2}\omega_B (p_{A2} - p_{B2})^2 p_{A1} q_{A1} p_{A2} q_{A2} p_{B1} q_{B1} (1 - \rho_A^2) \\ + (p_{A1} - p_{B1})(p_{A2} - p_{B2}) p_{A1} q_{A1} p_{A2} q_{A2} \sqrt{p_{B1} q_{B1} p_{B2} q_{B2}} \rho_B \omega_B (1 - \rho_A^2)$$

$$e = -\omega_B (p_{A1} - p_{B1})^2 p_{A1} q_{A1} p_{A2}^2 q_{A2}^2 (1 - \rho_A^2) \\ - \omega_B (p_{A2} - p_{B2})^2 p_{A1}^2 q_{A1}^2 p_{A2} q_{A2} (1 - \rho_A^2) \\ + 2(p_{A1} - p_{B1})(p_{A2} - p_{B2})(p_{A1} q_{A1} p_{A2} q_{A2})^{\frac{3}{2}} \rho_A \omega_B (1 - \rho_A^2).$$

Remark 2: One special case is that when $p_{A1} = p_{B1}$, $p_{A2} = p_{B2}$, $\rho_A = \rho_B = 0$, the expected weighted number of failures is:

$$w_1 q_{A1} n_A + w_1 q_{B1} n_B + w_2 q_{A2} n_A + w_2 q_{B2} n_B = w_2 (q_{A2} n_A + q_{B2} n_B) + w_1 q_{A1} n$$

We only need to minimize $q_{A2} n_A + q_{B2} n_B$, then $\omega = (q_{A2}, q_{B2})'$. We have

$$r = \max(\sqrt{p_{A2}} / \sqrt{p_{B2}}, \tilde{D}).$$

The above special case reduces to the optimal allocation proposed by Rosenberger *et al.* (2001), when there is no difference between the two groups for one endpoint and no

correlation between the endpoints. This is because the optimal allocations only depend on one endpoint.

Remark 3: Although the theoretical solutions of the optimal allocations are complicated, we can easily obtain these results with a low computing cost since they are functions of the estimated parameters.

3. Implementing the optimal allocation using response-adaptive randomization

In the literature, a variety of randomization approaches such as the sequential maximum likelihood (SMLE) (Melfi *et al.*, 2001) have been proposed for optimal allocation procedures. We implement the optimal allocations in Theorems 1 and 2 using the DBCD proposed by Hu and Zhang (2004), since its advantages are widely accepted. Assuming that the responses are available immediately, the DBCD can be described as follows:

1. We assign the first $2a$ subjects (usually 10%–20% of the total sample size) to the two treatment groups by certain restricted randomization procedures such as permuted block designs (Rosenberger and Lachin, 2002).
2. After $j-1$ ($j > 2a$) subjects have been assigned, we estimate the unknown parameters based on the $j-1$ observations, and we calculate the estimated optimal allocations based on Theorem 1 or 2.
3. We assign the j th subject to treatment A with probability ψ ,

$$\psi(v_A, \hat{\eta}_A^*) = \begin{cases} 1 & \text{if } v_A = 0 \\ \frac{\hat{\eta}_A^* (\hat{\eta}_A^* / v_A)^\gamma}{\hat{\eta}_A^* (\hat{\eta}_A^* / v_A)^\gamma + (1 - \hat{\eta}_A^*) ((1 - \hat{\eta}_A^*) / (1 - v_A))^\gamma} & \text{if } 0 < v_A < 1 \\ 0 & \text{if } v_A = 1 \end{cases}, \quad (4)$$

where v_A is the currently observed allocation proportion of the patients in treatment A, $\hat{\eta}_A^*$ is the estimated target allocation proportion, and γ is a tuning parameter that adjusts the balance between the extent of randomization and the convergence rates of the allocation proportion to the targets. In our numerical studies, the first 10% of the patients are allocated to the two treatments using restricted randomization to obtain initial parameter estimates, and then we switch to the DBCD procedure with $\gamma = 2$. There are few theoretical studies about the choice of the tuning parameters in the literature, though they are popular for adaptive randomization designs. Numerical studies showed that $\gamma = 2$ is a good trade-off (Rosenberger and Hu, 2004), and this value has been widely used in the literature (Tymofeyev *et al.*, 2007; Zhu and Hu, 2010, 2012). All the results are based on 5000 replications.

3.1. Simulation studies

First we study the type I error rate with $p_{A1} = p_{B1} = p_{A2} = p_{B2} = 0.3$. 500 patients are sequentially enrolled in the trial. Note that when the optimal allocation proportions in theorem 2 are the target, weights play an important role, whereas the optimal allocation proportions in Theorem 1 do not involve weights. We assign 7 different sets of weights for the scenario of Theorem 2. The relative weights (w_1/w_2) range from 0.1 to 10, and the corresponding importance of endpoint 2 decreases as the relative weights increase. In Table 1 we report the type I error rates (α) for different correlations between the two endpoints when DBCD is implemented targeting the optimal allocation proportions obtained in Theorems 1 and 2. We can see that type I error rate can be well controlled under both scenarios. In addition, the average parameter estimates and the corresponding standard deviations out of 5000 replications are also reported. Even if the RAR design involves complicated sequential updating procedures, we can still estimate the parameters very well.

We consider the following scenarios to study the efficient and ethical advantages of RAR procedures targeting the derived optimal allocation proportions in this paper:

Scenario 1: Treatment B is better than treatment A in terms of both endpoints, and the success rates are $p_{A1} = 0.20$, $p_{B1} = 0.35$, $p_{A2} = 0.30$, and $p_{B2} = 0.45$.

Scenario 2: The treatment effects of the two groups are the same in terms of endpoint 1, and treatment A is better than treatment B in terms of endpoint 2. The success rates are $p_{A1} = 0.25$, $p_{B1} = 0.25$, $p_{A2} = 0.45$, and $p_{B2} = 0.30$.

The above parameters are chosen to show the advantages of our procedure when the treatment effects are relatively conservative. RAR designs targeting the optimal allocation proportions may perform better for other parameter combinations.

In Table 2, we compare complete randomization (CR) with DBCD targeting the optimal allocation proportion obtained in Theorem 1, where power is the focus. The sample sizes (N) are chosen in each case to make the power to be about 0.8. We can see that our method consistently increases the power, which could lead to a reduction in the sample size in real trials.

In Table 3, we compare CR with DBCD targeting the optimal allocation proportion obtained in Theorem 2, and we report the EWNF and the power. Same weights as in Table 1 are assigned to those two endpoints. Table 3 shows that our procedures reduced the EWNF for all the cases while maintaining the same level of power. The advantages can be seen even in Scenario 2. Both Tables 2 and 3 show that the correlations have a dramatic influence on the sample size needed to achieve a given level of power.

Based on the asymptotic properties derived by Hu and Zhang (2004), the type I error rate for the DBCD can be easily controlled, as shown in Table 1. But it is worth studying the performance of our procedure in terms of type I error rate when the sample size is small. In Table 4, we report results for certain sample sizes in Table 2 when our procedure is targeting the optimal allocation proportions in Theorem 1 and $p_{A1} = p_{B1} = p_{A2} = p_{B2} = 0.3$. We find that the type I error rate is not quite satisfactory when the sample size is less than 250. This could be due to the convergence rate of the RAR procedures. The results are quite similar

when our procedure is targeting the optimal allocation proportions in Theorem 2 and the sample sizes are those in Table 3.

3.2. Re-designing the trial of perioperative total parenteral nutrition (TPN) treatment using the RAR design

Tang *et al.* (1989) discussed a trial comparing perioperative TPN treatment with no perioperative TPN in nutritionally compromised patients about to undergo surgery for gastric cancer. The investigators were interested in whether the perioperative TPN reduced the proportion of major and minor complications in the week following surgery. Each patient was at risk of major and minor complications that were well defined and mutually exclusive. Hence, these two endpoints were naturally correlated. Tang *et al.* (1989) carried out a numerical study with the following parameter combinations: $p_{A1} = 0.25$, $p_{B1} = 0.15$, $p_{A2} = 0.45$, and $p_{B2} = 0.30$. We redesign the trial and generate datasets using the same parameters. In Table 5, we report the results of DBCD targeting the optimal allocation obtained from Theorem 1 that aims to maximize the power, and we compare this procedure with CR. In Table 6, we report the results of DBCD targeting the optimal allocation obtained from Theorem 2 that aims to minimize the EWNF while maintaining the power, and we compare it with CR. In each scenario, we consider the correlations between the endpoints to range from 0 to 0.5, and we adjust the total sample size to achieve 80% power for CR. Tables 5 and 6 show that the RAR design can either achieve higher power of detecting the treatment effects under different levels of correlation between the endpoints, or reduce the number of failures while holding the power to the same level as that for CR.

4. Discussion

Multiple endpoints are widely employed in modern clinical trials. For patients with coronary heart disease, we may be interested in both resting and exercise ejection fractions. In blood-pressure trials, we might be interested in diastolic and systolic blood pressure or mean arterial pressure and pulse pressure. In stroke treatment, a number of scales are used to measure recovery, and no one scale is believed to assess all the dimensions. However, there has been little study of how to ensure that such trials are ethical and efficient through well-designed randomization procedures. In this paper, we use RAR designs to implement optimal allocation in order to maximize the power and minimize the EWNF. We give the analytical solutions to two general optimization problems for nonlinear models by using the KKT conditions. We have discussed the theoretical results and the advantages of our method in the previous sections. We now discuss possible concerns and directions for future research.

Dmitrienko et al. (2009) discussed different clinical decision rules and inferential goals. This paper focuses on how to achieve different ethical and efficient aims when a trial with multiple endpoints would be declared positive if at least one endpoint was significantly improved. In the Introduction section, we introduced the three-step approach for the development of optimal RAR procedures formalized by Hu and Rosenberger (2003). If other decision rules are employed and the objectives are changed, the corresponding optimization

problems such as equation (2) will be changed and the theoretically optimal allocation proportions will usually be different.

The trade-off between the power and the EWNF has long been discussed. We show that there exist procedures where the power is not compromised when more people are assigned to the better treatment. In general, adaptive designs may need a larger sample size than that for CR because of the larger variation in the allocation proportions. However, the traditional formula for the CR sample size also ignores the variation of the randomization procedure, so the actual loss of power under adaptive designs is less significant. This paper shows that targeting the appropriate allocation proportion with RAR designs can often increase the power and reduce the required sample size. The impression that equal allocation results in the best power is not always correct.

We also see that the extent of improvement of RAR over traditional clinical trial designs depends on lots of factors such as the objectives and parameter settings. In some cases, the improvement is not quite significant. But it is also worth noting that any improvement is desirable for certain deadly diseases such as Ebola. One of the major contributions of the paper is to derive the analytic optimal allocation for clinical trials with two possibly correlated endpoints. It is not only useful to understand the procedure but also provides foundation for future theoretical study and comparison.

The promotion of RAR procedures in real clinical trials require extensive research on its application in diverse practical scenarios. First, clinical trials with more than two endpoints or more than two treatments could be explored using a similar framework. On the one hand, the benefits of RAR procedures in multiarm clinical trials would be more impressive. On the other hand, more endpoints or treatments may require more unknown parameters. Second, other response types or a mixed type of endpoints such as one continuous and one binary endpoint could also be investigated. Third, when there are primary hypotheses and secondary hypotheses and the secondary hypotheses will be tested only when the primary hypotheses are rejected, the graphical approach to sequential testing (Bretz et al. 2009) is a useful tool for fixed design. When RAR procedures are used in this case, the graphical approach is potentially helpful to understand and define the objective of the problem and write out the objective function in the optimization problem. Fourth, Senn and Bretz (2007) discussed different types of powers, and innovative approaches such as using a latent variable were proposed. When RAR procedures are used in the above scenarios, we need to write out different objective or utility functions in the optimality problems. It may lead to difficulty in deriving the closed form of the optimal allocation proportions. However, we can still obtain the numerical solutions for the target allocations and implement them with appropriate randomization procedures. Since the RAR design targets the theoretically derived optimal allocation proportions, it usually performs better than the balanced complete randomization in terms of the targeted objective, unless the balanced allocation proportion is the theoretically optimal one. Moreover, we can also use RAR procedures to achieve some desirable features such as assigning more patients to the better treatment even if they are not from formal optimality problems. The control of type I error rate could be addressed using the theoretical results from Hu and Zhang (2004) and the innovative methods proposed for the scenario of interest such as Bretz et al. (2009) and Senn and Bretz (2007). Finally,

responses are assumed to be available immediately in this paper. In certain real clinical trials where the responses become available after most patients have been recruited and randomized, the RAR procedures are not appropriate in general. However, there is no logical difficulty applying RAR procedures in trials with moderate delayed response, given that some responses become available during the randomization period. We leave these topics for future research.

Acknowledgments

Research is supported by Hongjian Zhu's seeding funding from The University of Texas Health Science Center School of Public Health at Houston, and Yong Chen's NIH grant R21-LM012197, R01-LM009012 and R01-AI116794.

References

1. Biswas A, Bhattacharya R. Optimal response-adaptive designs for normal responses. *Biometrical Journal*. 2009; 51(1):193–202. [PubMed: 19197961]
2. Biswas A, Bhattacharya R. An optimal response-adaptive design with dual constraints. *Statistics & Probability Letters*. 2010; 80(3):177–185.
3. Biswas, A., Mandal, S. *mODa 7 — Advances in Model-Oriented Design and Analysis*. Springer; 2004. Optimal adaptive designs in phase III clinical trials for continuous responses with covariates; p. 51–59.
4. Bretz F, Maurer W, Brannath W, Posch M. A graphical approach to sequentially rejective multiple test procedures. *Statistics in medicine*. 2009; 28(4):586–604. [PubMed: 19051220]
5. Dmitrienko, A., Tamhane, AC., Bretz, F. *Multiple testing problems in pharmaceutical statistics*. CRC Press; 2009.
6. Friedman, LM., Furberg, C., DeMets, DL., Reboussin, DM., Granger, CB. *Fundamentals of clinical trials*. Vol. 4. Springer; 2010.
7. Gwise TE, Hu J, Hu F. Optimal biased coins for two-arm clinical trials. *Stat Interface*. 2008; 1:125–136.
8. Hu F, Rosenberger WF. Optimality, variability, power: Evaluating response-adaptive randomization procedures for treatment comparisons. *Journal of the American Statistical Association*. 2003; 98(463):671–678.
9. Hu, F., Rosenberger, WF. *The theory of response-adaptive randomization in clinical trials*. Vol. 525. John Wiley & Sons; 2006.
10. Hu F, Rosenberger WF, Zhang LX. Asymptotically best response-adaptive randomization procedures. *Journal of Statistical Planning and Inference*. 2006; 136(6):1911–1922.
11. Hu F, Zhang LX. Asymptotic properties of doubly adaptive biased coin designs for multitreatment clinical trials. *Annals of Statistics*. 2004; 32(1):268–301.
12. Ivanova A. A play-the-winner-type urn design with reduced variability. *Metrika*. 2003; 58(1):1–13.
13. Ivanova A, Rosenberger WF. A comparison of urn designs for randomized clinical trials of $k > 2$ treatments. *Journal of Biopharmaceutical Statistics*. 2000; 10(1):93–107. [PubMed: 10709804]
14. Ivanova A, Rosenberger WF. Adaptive designs for clinical trials with highly successful treatments. *Drug Information Journal*. 2001; 35(4):1087–1093.
15. Jennison C, Turnbull BW. Group sequential tests for bivariate response: Interim analyses of clinical trials with both efficacy and safety endpoints. *Biometrics*. 1993:741–752. [PubMed: 8241370]
16. Jennison, C., Turnbull, BW. *Group sequential methods with applications to clinical trials*. CRC Press; 1999.
17. Jeon Y, Hu F. Optimal adaptive designs for binary response trials with three treatments. *Statistics in Biopharmaceutical Research*. 2012
18. Karush, W. Master's thesis. Dept. of Mathematics, Univ. of Chicago; 1939. Minima of functions of several variables with inequalities as side constraints.

19. Kuhn, HW., Tucker, AW. Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability. University of California Press; Berkeley, Calif: 1951. Nonlinear programming; p. 481-492.
20. Melfi VF, Page C, Geraldes M. An adaptive randomized design with application to estimation. *Canadian Journal of Statistics*. 2001; 29(1):107-116.
21. Moyé, LA. Multiple analyses in clinical trials: fundamentals for investigators. Springer Science & Business Media; 2003.
22. Nocturnal Oxygen Therapy Trial Group. Continuous or nocturnal oxygen therapy in hypoxemic chronic obstructive lung disease: A clinical trial. *Ann Intern Med*. 1980; 93(3):391-398. [PubMed: 6776858]
23. Patnaik P. The non-central χ^2 - and f-distribution and their applications. *Biometrika*. 1949; 36(1/2): 202-232. [PubMed: 18132092]
24. Robbins H. Some aspects of the sequential design of experiments. *Bull Amer Math Soc*. 1952; 58(5):527-535.
25. Roberts R, Croft C, Gold HK, Hartwell TD, Jaffe AS, Muller JE, Mullin SM, Parker C, Passamani ER, Poole WK, et al. Effect of propranolol on myocardial-infarct size in a randomized blinded multicenter trial. *New England Journal of Medicine*. 1984; 311(4):218-225. [PubMed: 6377070]
26. Rosenberger WF, Hu F. Maximizing power and minimizing treatment failures in clinical trials. *Clinical Trials*. 2004; 1(2):141-147. [PubMed: 16281886]
27. Rosenberger WF, Lachin JM. Randomization and the clinical trial. *Randomization in Clinical Trials: Theory and Practice*. 2002:1-14.
28. Rosenberger WF, Stallard N, Ivanova A, Harper CN, Ricks ML. Optimal adaptive designs for binary response trials. *Biometrics*. 2001; 57(3):909-913. [PubMed: 11550944]
29. Senn S, Bretz F. Power and sample size when multiple endpoints are considered. *Pharmaceutical Statistics*. 2007; 6(3):161-170. [PubMed: 17674404]
30. Tang DI, Gnecco C, Geller NL. Design of group sequential clinical trials with multiple endpoints. *Journal of the American Statistical Association*. 1989; 84(407):775-779.
31. Thompson WR. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*. 1933; 25(3-4):285-294.
32. Tymofyeyev Y, Rosenberger WF, Hu F. Implementing optimal allocation in sequential binary response experiments. *Journal of the American Statistical Association*. 2007; 102(477):224-234.
33. Urokinase Pulmonary Embolism Trial Study Group. Urokinasestreptokinase embolism trial: Phase 2 results. *Journal of the American Medical Association*. 1974; 229:1606-1613. [PubMed: 4408392]
34. Wei L, Durham S. The randomized play-the-winner rule in medical trials. *Journal of the American Statistical Association*. 1978; 73(364):840-843.
35. Zeiher BG, Artigas A, Vincent JL, Dmitrienko A, Jackson K, Thompson BT, Bernard G, Group SS, et al. Neutrophil elastase inhibition in acute lung injury: results of the strive study. *Critical care medicine*. 2004; 32(8):1695-1702. [PubMed: 15286546]
36. Zelen M. Play the winner rule and the controlled clinical trial. *Journal of the American Statistical Association*. 1969; 64(325):131-146.
37. Zhang L, Rosenberger WF. Response-adaptive randomization for clinical trials with continuous outcomes. *Biometrics*. 2006; 62(2):562-569. [PubMed: 16918921]
38. Zhang L, Rosenberger WF. Response-adaptive randomization for survival trials: The parametric approach. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*. 2007; 56(2):153-165.
39. Zhu H, Hu F. Sequential monitoring of response-adaptive randomized clinical trials. *The Annals of Statistics*. 2010; 38(4):2218-2241.
40. Zhu H, Hu F. Interim analysis of clinical trials based on urn models. *Canadian Journal of Statistics*. 2012; 40(3):550-568.

APPENDIX: PROOFS

Proof of Theorem 1

This proof involves lengthy derivations and calculations; we emphasize the major ideas and steps. To prove the main theorem we need the following lemma.

Lemma 1

The noncentrality parameter $\phi(n_A, n_B)$ in (1) is a concave function, and $\nabla \phi \geq 0$.

We omit the proof of the concaveness of ϕ because it is similar to the proof in Tymofeyev *et al.* (2007); we show that $\nabla \phi \geq 0$.

Let $\Delta_1 = p_{A1} - p_{B1}$, $\Delta_2 = p_{A2} - p_{B2}$, $A_1 = p_{A1}q_{A1}$, $A_2 = p_{A2}q_{A2}$, $B_1 = p_{B1}q_{B1}$, and $B_2 = p_{B2}q_{B2}$. Then

$$\phi(\mathbf{n}) = \left\{ \Delta_1^2 \left(\frac{A_2}{n_A} + \frac{B_2}{n_B} \right) + \Delta_2^2 \left(\frac{A_1}{n_A} + \frac{B_1}{n_B} \right) - 2\psi \Delta_1 \Delta_2 \right\} / \left\{ \left(\frac{A_1}{n_A} + \frac{B_1}{n_B} \right) \left(\frac{A_2}{n_A} + \frac{B_2}{n_B} \right) - \psi^2 \right\} = \frac{a(\mathbf{n})}{S(\mathbf{n})}$$

where $\psi = \rho_A \sqrt{A_1 A_2 / n_A} + \rho_B \sqrt{B_1 B_2 / n_B}$.

$$\begin{aligned} \frac{\partial \psi}{\partial n_K} &= -\rho_K \frac{\sqrt{K_1 K_2}}{n_K^2}, \quad K=A, B \\ \frac{\partial a(\mathbf{n})}{\partial n_K} &= -\frac{\Delta_1^2 K_2}{n_K^2} - \frac{\Delta_2^2 K_1}{n_K^2} + 2\Delta_1 \Delta_2 \rho_K \frac{\sqrt{K_1 K_2}}{n_K^2}, \quad K=A, B \\ \frac{\partial S(\mathbf{n})}{\partial n_K} &= -\frac{K_1}{n_K^2} \left(\frac{A_2}{n_A} + \frac{B_2}{n_B} \right) - \frac{K_2}{n_K^2} \left(\frac{A_1}{n_A} + \frac{B_1}{n_B} \right) + 2\psi \rho_K \frac{\sqrt{K_1 K_2}}{n_K^2}, \quad K=A, B \\ \frac{\partial \phi(\mathbf{n})}{\partial n_K} &= \frac{1}{S(\mathbf{n})^2} \left\{ \frac{\partial a(\mathbf{n})}{\partial n_K} S(\mathbf{n}) - a(\mathbf{n}) \frac{\partial S(\mathbf{n})}{\partial n_K} \right\} \\ &= \left[\left\{ K_2 \psi^2 + K_1 \left(\frac{A_2}{n_A} + \frac{B_2}{n_B} \right)^2 - 2\sqrt{K_1 K_2} \left(\frac{A_1}{n_A} + \frac{B_1}{n_B} \right) \psi \rho_K \right\} M^2 \right. \\ &\quad \left. + K_1 K_2 \Delta_2^2 \left\{ \psi^2 - \left(\frac{A_1}{n_A} + \frac{B_1}{n_B} \right) \left(\frac{A_2}{n_A} + \frac{B_2}{n_B} \right) \right\}^2 (1 - \rho_K^2) \right] / S(\mathbf{n})^2 n_K^2 \end{aligned}$$

$$\begin{aligned} M &= \Delta_1 - K_2 \psi \Delta_2 \left(A_1 / n_A \right. \\ &\quad \left. + B_1 / n_B \right) \\ &\quad - K_1 \psi \Delta_2 \left(A_2 / n_A \right. \\ &\quad \left. + B_2 / n_B \right) + \sqrt{K_1 K_2} \psi^2 \Delta_2 + \sqrt{K_1 K_2} \Delta_2 \left(A_1 / n_A \right. \\ &\quad \left. + B_1 / n_B \right) \left(A_2 / n_A \right. \\ &\quad \left. + B_2 / n_B \right) \rho_K \end{aligned}$$

where $\rho_K = \rho_A \rho_B$, $K = A, B$.

We have that

$$\begin{aligned} &K_2 \psi^2 + K_1 \left(\frac{A_2}{n_A} + \frac{B_2}{n_B} \right)^2 - 2\sqrt{K_1 K_2} \left(\frac{A_1}{n_A} + \frac{B_1}{n_B} \right) \psi \rho_K \\ &= \left\{ \sqrt{K_2} \psi - \sqrt{K_1} \left(\frac{A_2}{n_A} + \frac{B_2}{n_B} \right) \right\}^2 + 2\sqrt{K_1 K_2} \left(\frac{A_1}{n_A} + \frac{B_1}{n_B} \right) \psi (1 - \rho_K) \geq 0 \end{aligned}$$

when $0 < \rho_K < 1$.

Thus, we have $\nabla \phi = 0$.

For this convex problem, the first-order Karush-Kuhn-Tucker (KKT) necessary conditions for optimality become sufficient conditions. We need to find the Lagrange multipliers and the optimal allocation $\mathbf{n}^*$ satisfying the following KKT conditions:

$$\nabla \phi(\mathbf{n}^*) - \lambda \mathbf{w} + \lambda_A (\mathbf{e}_1 - D\mathbf{1}) + \lambda_B (\mathbf{e}_2 - D\mathbf{1}) = 0 \quad (\text{A.1})$$

$$\lambda (\mathbf{w}' \mathbf{n}^* - \mathbf{N}) = 0 \quad (\text{A.2})$$

$$\lambda_K (n_K^* - D(n_A^* + n_B^*)) = 0, \quad K = A, B \quad (\text{A.3})$$

$$\lambda \geq 0, \lambda_K \geq 0, \quad K = A, B \quad (\text{A.4})$$

where $\mathbf{e}_j$ is a vector of all 0's except for a one in the j th position, $j = 1, 2$.

In Theorem 1 we let $\mathbf{w} = (1, 1)'$ and write condition (A.1) in terms of $\eta_A^* = n_A^* / (n_A^* + n_B^*)$. Condition (A.2) can always be satisfied by changing the scaling of the η_A^* 's. Then we have

$$\lambda_K (\eta_K^* - D) = 0, \quad K = A, B \quad (\text{A.5})$$

$$-\nabla \phi(\boldsymbol{\eta}^*) + \lambda \mathbf{1} + \lambda_A \begin{pmatrix} D-1 \\ D \end{pmatrix} + \lambda_B \begin{pmatrix} D \\ D-1 \end{pmatrix} = 0. \quad (\text{A.6})$$

When $\eta_A^* \neq 0, \eta_B^* \neq 0$, we substitute $\boldsymbol{\eta}^*$ into (A.2) to get:

$$\begin{aligned}
& \left[\eta_B^4 \left\{ -A_1 A_2^2 \Delta_1^2 (\rho_A^2 - 1) - A_1^2 A_2 \Delta_2^2 (\rho_A^2 - 1) + 2 A_1 A_2 \sqrt{A_1 A_2} \Delta_1 \Delta_2 \rho_A (\rho_A^2 - 1) \right\} \right. \\
& + \eta_A \eta_B^3 \left\{ -2 A_1 A_2 B_2 \Delta_1^2 (\rho_A^2 - 1) - 2 A_1 A_2 B_1 \Delta_2^2 (\rho_A^2 - 1) + 4 A_1 A_2 \sqrt{B_1 B_2} \Delta_1 \Delta_2 \right. \\
& \quad \times (\rho_A^2 - 1) \rho_B \} + \eta_A^2 \eta_B \left\{ 2 A_2 B_1 B_2 \Delta_1^2 (\rho_B^2 - 1) + 2 A_1 B_1 B_2 \Delta_2^2 (\rho_B^2 - 1) \right. \\
& \quad \left. \left. - 4 \sqrt{A_1 A_2} B_1 B_2 \Delta_1 \Delta_2 \rho_A (\rho_B^2 - 1) \right\} \right. \\
& + \eta_A^4 \left\{ B_1 B_2^2 \Delta_1^2 (\rho_B^2 - 1) + B_1^2 B_2 \Delta_2^2 (\rho_B^2 - 1) - 2 B_1 B_2 \sqrt{B_1 B_2} \Delta_1 \Delta_2 \rho_B (\rho_B^2 - 1) \right\} \\
& + \eta_A^2 \eta_B^2 \left\{ -A_2^2 B_1 \Delta_1^2 + A_1 B_2^2 \Delta_1^2 + A_2 B_1^2 \Delta_2^2 - A_1^2 B_2 \Delta_2^2 - A_1 A_2 B_2 \Delta_1^2 \rho_A^2 \right. \\
& \quad - A_1 A_2 B_1 \Delta_2^2 \rho_B^2 - 2 \sqrt{A_1 A_2} \sqrt{B_1 B_2} (-A_2 + B_2) \Delta_1^2 \rho_A \rho_B \\
& \quad + 2 A_1 \sqrt{A_1 A_2} \sqrt{B_1 B_2} \Delta_2^2 \rho_A \rho_B - 2 \sqrt{A_1 A_2} B_1 \sqrt{B_1 B_2} \Delta_2^2 \rho_A \rho_B \\
& \quad - 2 A_1 A_2 \sqrt{B_1 B_2} \Delta_1 \Delta_2 (1 + \rho_A^2) \rho_B + A_2 B_1 B_2 \Delta_1^2 \rho_B^2 + A_1 B_1 B_2 \Delta_2^2 \rho_B^2 \\
& \quad \left. - 2 A_2 B_1 \Delta_1 \Delta_2 (-\sqrt{A_1 A_2} \rho_A + \sqrt{B_1 B_2} \rho_B) - 2 A_1 B_2 \Delta_1 \Delta_2 (-\sqrt{A_1 A_2} \rho_A \right. \\
& \quad \left. + \sqrt{B_1 B_2} \rho_B) + 2 \sqrt{A_1 A_2} B_1 B_2 \Delta_1 \Delta_2 \rho_A (1 + \rho_B^2) \right\} \Big] \\
& \quad \times \left[\eta_B \left\{ -A_1 (B_2 \eta_A + A_2 \eta_B - A_2 \eta_B \rho_A^2) + 2 \sqrt{A_1 A_2} \sqrt{B_1 B_2} \eta_A \rho_A \rho_B \right\} \right. \\
& \quad \left. + B_1 \eta_A \left\{ -A_2 \eta_B + B_2 \eta_A (\rho_B^2 - 1) \right\} \right]^{-2} = 0. \tag{*}
\end{aligned}$$

Setting $r = \eta_A/\eta_B$, with a simple calculation (*) becomes a quartic function: $ar_4 + br_3 + cr_2 + dr + e = 0$, where a, b, c, d , and e are as given in Theorem 1. We solve this quartic function and choose the roots between 0 and 1 to get the solutions.

For Theorem 2, we let $\mathbf{w} = (w_1 q_{A1} + w_2 q_{A2}, w_1 q_{B1} + w_2 q_{B2})'$ and obtain the solutions similarly. We use Mathematica to help derive the theoretically optimal allocations in Theorems 1 and 2.

- Clinical trials with multiple endpoints were studied.
- Optimal allocation proportions for maximizing the power of the test of homogeneity with a fixed sample size have been derived.
- Optimal allocation proportions for minimizing the expected weighted number of failures with a fixed power have been derived.
- Response-adaptive randomization procedures were used to implement these optimal allocations and efficient and ethical advantages have been shown.

Table 1
Type I error rates and parameter estimates for DBCD targeting optimal allocations in Theorems 1 and 2

ρ_A	ρ_B	w_1	w_2	$\hat{p}_{A1}$	$\hat{p}_{A2}$	$\hat{p}_{B1}$	$\hat{p}_{B2}$	$\hat{p}_A$	$\hat{p}_B$	α
Targeting optimal allocations in Theorem 2										
0	0	1	0.1	0.300(0.029)	0.299(0.029)	0.300(0.029)	0.299(0.029)	0.001(0.063)	0.001(0.063)	0.051
0	0	1	0.25	0.299(0.029)	0.300(0.029)	0.300(0.029)	0.300(0.030)	0.000(0.063)	0.002(0.063)	0.052
0	0	1	0.5	0.300(0.029)	0.299(0.029)	0.299(0.029)	0.300(0.029)	0.000(0.063)	0.000(0.064)	0.053
0	0	1	1	0.299(0.029)	0.300(0.029)	0.299(0.029)	0.299(0.029)	0.000(0.063)	0.000(0.063)	0.051
0	0	1	2	0.300(0.029)	0.300(0.029)	0.300(0.029)	0.299(0.029)	-0.001(0.063)	-0.001(0.063)	0.051
0	0	1	4	0.300(0.029)	0.300(0.029)	0.300(0.029)	0.299(0.029)	0.000(0.063)	-0.001(0.063)	0.053
0	0	1	10	0.300(0.029)	0.300(0.029)	0.300(0.029)	0.299(0.029)	0.000(0.063)	0.001(0.063)	0.049
0.3	0.3	1	0.1	0.300(0.029)	0.299(0.029)	0.300(0.029)	0.299(0.029)	0.300(0.065)	0.300(0.065)	0.049
0.3	0.3	1	0.25	0.299(0.028)	0.300(0.029)	0.300(0.029)	0.300(0.030)	0.300(0.065)	0.301(0.065)	0.053
0.3	0.3	1	0.5	0.299(0.029)	0.300(0.029)	0.300(0.029)	0.300(0.029)	0.300(0.065)	0.301(0.066)	0.049
0.3	0.3	1	1	0.299(0.028)	0.300(0.029)	0.299(0.029)	0.299(0.029)	0.301(0.065)	0.300(0.066)	0.052
0.3	0.3	1	2	0.300(0.029)	0.300(0.029)	0.300(0.028)	0.299(0.029)	0.299(0.064)	0.298(0.065)	0.052
0.3	0.3	1	4	0.300(0.029)	0.300(0.029)	0.300(0.028)	0.299(0.029)	0.299(0.063)	0.298(0.063)	0.052
0.3	0.3	1	10	0.300(0.029)	0.300(0.029)	0.300(0.029)	0.299(0.029)	0.298(0.064)	0.301(0.066)	0.052
0.5	0.3	1	0.1	0.299(0.029)	0.299(0.029)	0.300(0.029)	0.299(0.029)	0.501(0.060)	0.300(0.065)	0.050
0.5	0.3	1	0.25	0.300(0.029)	0.300(0.029)	0.299(0.029)	0.300(0.030)	0.499(0.061)	0.300(0.065)	0.053
0.5	0.3	1	0.5	0.299(0.029)	0.300(0.029)	0.299(0.029)	0.300(0.029)	0.500(0.061)	0.301(0.065)	0.050
0.5	0.3	1	1	0.300(0.029)	0.299(0.029)	0.299(0.029)	0.299(0.029)	0.500(0.060)	0.300(0.065)	0.050
0.5	0.3	1	2	0.299(0.029)	0.300(0.029)	0.300(0.028)	0.299(0.029)	0.499(0.060)	0.298(0.065)	0.049
0.5	0.3	1	4	0.300(0.029)	0.300(0.029)	0.300(0.029)	0.299(0.029)	0.500(0.059)	0.299(0.064)	0.048
0.5	0.3	1	10	0.299(0.029)	0.300(0.029)	0.300(0.029)	0.299(0.029)	0.499(0.060)	0.300(0.065)	0.049
Targeting optimal allocations in Theorem 1										
0	0	-	-	0.299(0.029)	0.299(0.028)	0.300(0.030)	0.299(0.028)	0.001(0.063)	0.001(0.062)	0.050
0.3	0.3	-	-	0.299(0.030)	0.299(0.029)	0.300(0.030)	0.299(0.028)	0.300(0.066)	0.300(0.065)	0.053
0.5	0.3	-	-	0.299(0.030)	0.299(0.029)	0.300(0.030)	0.299(0.028)	0.500(0.061)	0.300(0.064)	0.052

Table 2

Comparison of complete randomization and DBCD targeting the optimal allocation in Theorem 1

ρ_A	ρ_B	N	Power	
			CR	RAR
$p_{A1} = 0.20$ $p_{B1} = 0.35$ $p_{A2} = 0.30$ $p_{B2} = 0.45$				
0.0	0.0	178	0.794	0.808
0.1	0.1	197	0.798	0.811
0.3	0.1	210	0.800	0.809
0.3	0.3	234	0.803	0.809
0.5	0.3	250	0.799	0.805
$p_{A1} = 0.25$ $p_{B1} = 0.25$ $p_{A2} = 0.45$ $p_{B2} = 0.30$				
0.0	0.0	395	0.798	0.809
0.1	0.1	395	0.802	0.807
0.3	0.1	380	0.800	0.811
0.3	0.3	360	0.801	0.807
0.5	0.3	334	0.798	0.811

Table 3

Comparison of complete randomization and DBCD targeting the optimal allocation in Theorem 2 in terms of EWNF* and power

$p_{A1} = 0.20$ $p_{B1} = 0.35$ $p_{A2} = 0.30$ $p_{B2} = 0.45$												
ρ_A	ρ_B	N	w_1	1	1	1	1	1	1	1	1	1
			w_2	0.1	0.25	0.5	1	2	4	10		
0	0	178	CR	140	157	185	241	352	575	1244		
			RAR	139	155	183	238	348	568	1228		
			Power	0.803	0.801	0.803	0.806	0.797	0.806	0.802	(0.794 **)	
0.1	0.1	197	CR	155	174	204	266	390	636	1377		
			RAR	153	171	201	262	383	625	1349		
			Power	0.802	0.809	0.803	0.794	0.805	0.802	0.808	(0.798)	
0.3	0.1	210	CR	165	185	218	284	415	678	1467		
			RAR	164	183	215	280	410	669	1444		
			Power	0.803	0.803	0.809	0.805	0.807	0.803	0.804	(0.800)	
0.3	0.3	234	CR	184	206	243	316	463	756	1634		
			RAR	182	203	239	311	455	742	1606		
			Power	0.804	0.806	0.802	0.802	0.805	0.808	0.805	(0.803)	
0.5	0.3	250	CR	197	221	260	338	494	807	1745		
			RAR	195	218	256	333	488	796	1722		
			Power	0.797	0.800	0.804	0.803	0.802	0.803	0.803	(0.799)	
$p_{A1} = 0.25$ $p_{B1} = 0.25$ $p_{A2} = 0.45$ $p_{B2} = 0.30$												
0	0	395	CR	322	359	420	544	791	1285	2768		
			RAR	320	357	419	541	785	1273	2735		
			Power	0.806	0.801	0.807	0.803	0.799	0.808	0.805	(0.798)	
0.1	0.1	395	CR	321	359	420	544	791	1285	2768		
			RAR	320	357	419	541	785	1273	2736		
			Power	0.805	0.801	0.802	0.799	0.803	0.806	0.802	(0.802)	
0.3	0.1	380	CR	309	345	405	524	761	1237	2665		
			RAR	308	344	403	521	756	1226	2637		
			Power	0.811	0.807	0.797	0.805	0.803	0.808	0.808	(0.800)	

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

$p_{A1} = 0.20$ $p_{B1} = 0.35$ $p_{A2} = 0.30$ $p_{B2} = 0.45$												
ρ_A	ρ_B	N	w_1	1	1	1	1	1	1	2	4	10
			w_2									
0.3	0.3	360	CR	293	327	383	496	721	1171	2522		
			RAR	292	325	381	493	716	1160	2496		
			Power	0.806	0.806	0.799	0.806	0.803	0.804	0.805	(0.801)	
0.5	0.3	334	CR	272	303	356	460	669	1087	2340		
			RAR	271	302	354	458	666	1080	2321		
			Power	0.807	0.806	0.803	0.796	0.807	0.796	0.805	(0.798)	

*. For each combination of ρ_A and ρ_B , the first line is EWNF of CR procedure, and the second line is EWNF of DBCD.

**. For each combination of ρ_A and ρ_B , the third line is the power of DBCD as well as the power of CR that is reported in parentheses for comparison.

Type I error rates and parameter estimates for DBCD targeting optimal allocations in Theorems 1 when sample sizes are smaller

Table 4

N	ρ_A	ρ_B	$\hat{p}_{A1}$	$\hat{p}_{A2}$	$\hat{p}_{B1}$	$\hat{p}_{B2}$	$\hat{\rho}_A$	$\hat{\rho}_B$	α
178	0	0	0.299(0.050)	0.299(0.048)	0.299(0.049)	0.300(0.048)	0.007(0.107)	0.008(0.103)	0.060
234	0.3	0.3	0.299(0.043)	0.299(0.042)	0.299(0.043)	0.298(0.042)	0.300(0.098)	0.300(0.095)	0.062
250	0.5	0.3	0.299(0.042)	0.299(0.041)	0.299(0.042)	0.298(0.040)	0.502(0.088)	0.298(0.091)	0.057
334	0.5	0.3	0.299(0.036)	0.299(0.035)	0.300(0.036)	0.299(0.034)	0.502(0.076)	0.300(0.078)	0.053
360	0.3	0.3	0.299(0.034)	0.299(0.034)	0.300(0.035)	0.299(0.033)	0.300(0.077)	0.301(0.077)	0.052
395	0	0	0.300(0.033)	0.300(0.032)	0.300(0.033)	0.299(0.032)	0.002(0.072)	0.002(0.069)	0.049

Table 5

Comparison of complete randomization and DBCD targeting the optimal allocation in Theorem 1 for redesigned trial

ρ_A	ρ_B	N	Power	
			CR	RAR
0	0	240	0.802	0.818
0.1	0.1	260	0.799	0.807
0.2	0.2	285	0.801	0.812
0.3	0.1	280	0.796	0.802
0.3	0.3	300	0.799	0.802
0.5	0.3	335	0.795	0.813
0.5	0.5	349	0.797	0.805

Table 6

Comparison of complete randomization and DBCD targeting the optimal allocation in Theorem 2 for redesigned trial in terms of EWNF* and power

ρ_A	ρ_B	N	w_1	1	1	1	1	1	1	1	1	1
0	0	240	w_2	0.1	0.25	0.5	1	1	2	4	10	
0.1	0.1	260	CR	207	229	266	342	492	792	1692		
			RAR	205	227	264	338	485	779	1665		
			Power	0.809	0.808	0.807	0.806	0.799	0.804	0.804	(0.802**)	
0.2	0.2	285	CR	225	248	289	370	533	858	1834		
			RAR	222	246	286	366	525	845	1802		
			Power	0.800	0.803	0.798	0.802	0.807	0.799	0.805	(0.799)	
0.3	0.3	240	CR	246	272	317	406	584	941	2010		
			RAR	243	269	313	401	575	925	1975		
			Power	0.802	0.808	0.810	0.810	0.804	0.802	0.806	(0.801)	
0.3	0.3	240	CR	242	267	311	399	574	924	1975		
			RAR	239	264	307	393	564	907	1935		
			Power	0.801	0.797	0.800	0.797	0.796	0.798	0.798	(0.796)	
0.5	0.3	240	CR	259	287	334	427	615	991	2118		
			RAR	256	283	329	421	605	973	2076		
			Power	0.798	0.800	0.800	0.797	0.801	0.801	0.798	(0.799)	
0.5	0.3	240	CR	290	320	372	477	687	1106	2364		
			RAR	286	316	367	470	674	1083	2308		
			Power	0.808	0.811	0.810	0.809	0.806	0.811	0.804	(0.795)	

* For each combination of ρ_A and ρ_B , the first line is EWNF of CR procedure, and the second line is EWNF of DBCD.

** For each combination of ρ_A and ρ_B , the third line is the power of DBCD as well as the power of CR that is reported in parentheses for comparison.