



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

ARCHIVIO ISTITUZIONALE  
DELLA RICERCA

## Alma Mater Studiorum Università di Bologna Archivio istituzionale della ricerca

Optimal designs for testing hypothesis in multiarm clinical trials

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

*Published Version:*

Alessandro Baldi Antognini, M.N. (2019). Optimal designs for testing hypothesis in multiarm clinical trials. STATISTICAL METHODS IN MEDICAL RESEARCH, 28(10-11), 3242-3259 [10.1177/0962280218797960].

*Availability:*

This version is available at: <https://hdl.handle.net/11585/657011> since: 2020-01-17

*Published:*

DOI: <http://doi.org/10.1177/0962280218797960>

*Terms of use:*

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).  
When citing, please refer to the published version.

(Article begins on next page)

This is the final peer-reviewed accepted manuscript of:

**Baldi Antognini A, Novelli M, Zagoraiou M. Optimal designs for testing hypothesis in multiarm clinical trials. *Statistical Methods in Medical Research*. 2019;28(10-11):3242-3259. doi:10.1177/0962280218797960**

The final published version is available online at:

<https://doi.org/10.1177/0962280218797960>

Rights / License:


The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

*This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>)*

***When citing, please refer to the published version.***

---

# Optimal designs for testing hypothesis in multiarm clinical trials

Journal Title  
XX(X):2–26  
©The Author(s) 0000  
Reprints and permission:  
sagepub.co.uk/journalsPermissions.nav  
DOI: 10.1177/ToBeAssigned  
www.sagepub.com/  


Alessandro Baldi Antognini<sup>1</sup>, Marco Novelli<sup>1</sup> and Maroussa Zagoraïou<sup>1</sup>

## Abstract

The present paper deals with the problem of designing randomized multiarm clinical trials for treatment comparisons in order to achieve a suitable trade-off among inferential precision and ethical concerns. Although the large majority of the literature is focused on the estimation of the treatment effects, in particular for the case of two treatments with binary outcomes, the present paper takes into account the inferential goal of maximizing the power of statistical tests to detect correct conclusions about the treatment effects for normally response trials. After discussing the allocation optimizing the power of the classical multivariate test of homogeneity, we suggest a multipurpose design methodology, based on constrained optimization, which maximizes the power of the test under a suitable ethical constraint reflecting the effectiveness of the treatments. The ensuing optimal allocation depends in general on the unknown model parameters but, contrary to the unconstrained optimal solution or to some targets proposed in the literature, it is a non-degenerate continuous function of the treatment contrasts and therefore it can be approached by standard response-adaptive randomization procedures. The properties of this constrained optimal allocation are described both theoretically and through suitable examples, showing good performances both in terms of ethical gain and statistical efficiency, taking into account estimation precision as well.

## Keywords

Asymptotic inference, Ethics, Power, Response-adaptive designs, Wald test.

## 1 Introduction

Starting from the well-known Efron's Biased Coin Design<sup>1</sup>, several randomized clinical trials have been designed in order to achieve balanced or nearly balanced allocations among the treatment groups, with the aim of maximizing inferential precision about the treatment comparisons. The main justification concerns the optimal properties of balance, which is often regarded as desirable from the viewpoint of estimation accuracy. Indeed, under the linear homoscedastic model setup, balance is universally optimal<sup>2</sup>, **i.e., the balanced design optimizes every design criterion applied to the variance-covariance matrix of the parameter estimators. In particular, it optimizes the usual criteria for the estimation of the treatment effects like, e.g., the well-known  $D$ -optimality (minimizing the volume of the confidence ellipsoid of the parameters of interest), or the  $tr$ -optimality (i.e., trace-optimality, also called  $A$ -optimality, minimizing the sum of the variances of the estimators).**

Although balancing the allocations is usually considered optimal for estimation under homoscedasticity, it is also appealing even in the case of heteroscedastic treatment groups, since it is still  $D$ -optimal for the joint estimation of the treatment effects and it is nearly optimal under several optimality criteria<sup>3,4</sup>.

Clearly, maximizing the power to detect a significant treatment difference is another major issue in designing clinical trials. In the case of two treatments, balance is optimal also from the viewpoint of testing, since it maximizes the power for homoscedastic outcomes<sup>5</sup> and it is asymptotically optimal for binary trials<sup>6</sup>. However, in the case of several treatments, the inferential goals could be focused on the treatment contrasts, in place of the main effects, and the balanced allocation may not be efficient. Although balance is still  $D$ -optimal for the joint estimation of the contrasts under homoscedasticity, it is not  $tr$ -optimal (**namely, balance does not minimize the average variance of the estimators of the treatment contrasts**) and it is significantly different from the optimal design for hypothesis testing in binary trials<sup>7</sup>. Moreover, balance could be strongly inappropriate in clinical trials in which ethical demands play a fundamental role, like **oncological trials or clinical experiments for rare diseases**, where the demand of individual care often induces to skew the allocations to more efficacious (or less toxic) treatments. Frequently, the

---

<sup>1</sup>Department of Statistical Sciences, University of Bologna, Italy.

**Corresponding author:**

Alessandro Baldi Antognini, Department of Statistical Sciences, University of Bologna, Via delle Belle Arti 41, 40126, Bologna, Italy.

Email: a.baldi@unibo.it

conflicting goals related to the ethical concern of maximizing the subjects care and to the statistical aim of drawing correct inferential conclusions with high precision can be formalized into suitable combined/constrained optimization problems and several authors derived target allocations that could be regarded as a valid trade-off among ethics and inference<sup>7-13</sup>. Generally, these allocations depend on the unknown parameters and, under suitable conditions, they can be targeted by using response-adaptive (RA) randomization procedures, namely sequential rules that change at each step the allocations probabilities of the treatments to approximate the chosen target (e.g., the doubly-adaptive biased coin design<sup>14</sup>). For a recent review see the books<sup>15-17</sup>.

However, most of the target allocations for multiarm clinical trials proposed in the literature are focused on estimation precision; while little attention has been devoted to hypothesis testing, almost exclusively for binary trials. In particular, Tymofyeyev et al.<sup>7</sup> derived the target maximizing the power of Wald test of homogeneity subject to i) an ethical constraint about the upper bound for the total expected number of failures and ii) a lower bound for each treatment allocation proportion. Although the subjective choice of these thresholds induces flexibility, the ensuing target is a discontinuous function of the unknown model parameters, whose closed-form expression is not generally available. Moreover, excluding binary trials, the suggested target allocations are quite few, often of heuristic nature or characterized by complicated functional forms<sup>13;18-20</sup>.

**The present paper deals with the problem of designing randomized multiarm clinical trials for treatment comparisons for the linear homoscedastic model. Although in the literature most attention has been focused on binary outcomes, there are many clinical trials where, especially in the design phase, the primary endpoint could be assumed normally distributed (like, e.g., the diastolic blood pressure reduction<sup>21</sup> or the duration of cold symptoms<sup>22</sup>). Furthermore, in several circumstances the outcomes can be treated as approximately normal either via suitable transformation of the data<sup>23</sup> or when the primary outcomes are individual means of repeated measurements/scores<sup>24</sup>). As a motivating example, Dworkin et al. (2003) discussed a randomized placebo-controlled trial intended to investigate the efficacy of pregabalin in the treatment of postherpetic neuralgia. The efficacy measure was a 11-point numerical pain rating score and the primary endpoint was the mean pain score derived from patients' last 7 days of diary entries while taking the study drug. According to several authors (see e.g.<sup>9;10;15;25</sup>), the observed mean scores in the two groups exhibit an approximate normal distribution with different means but substantially equal standard deviations.**

The first aim of the paper is to derive the allocation maximizing the power of the classical multivariate test about the treatment contrasts. This optimal target

is a degenerate allocation which depends on the unknown model parameters and assigns 50% of the subjects both to the best and the worst treatment. Therefore, we suggest a multipurpose design strategy that combines ethical demands and the inferential goal of maximizing the power of the statistical test. Adopting a constrained optimization framework, the suggested methodology is aimed at maximizing the non-centrality parameter of the classical Wald test of homogeneity subject to an ethical constraint on the treatment allocation proportions reflecting the efficacy of the competing drugs. When the ethical gain tends to be small, the ensuing constrained optimal target coincides with the balanced one; whereas, as the relative superiority of the best treatment (wrt the other drugs) increases, it skews the assignments to the superior treatment maintaining all the others equireplicated. This target admits a very simple analytical form and, contrary to the unconstrained optimal solution, it is a non-degenerate continuous function of the treatment contrasts, so it can be approached by standard response-adaptive randomization procedures. The properties of this constrained optimal allocation are described both theoretically and through suitable examples, showing a very good behavior wrt some traditional methods both in terms of ethical gain and statistical efficiency, taking into account estimation precision too.

The paper is structured as follows. In Section 2, the model and some preliminaries are presented, while Section 3 deals with optimal target allocations for hypothesis testing, taking into account both constrained and unconstrained optimization, as well as their properties. Section 4 discusses the performance of the proposed targets through numerical examples, also compared with those of other allocations suggested in the literature, **while Section 5 highlights some general conclusions**. Mathematical details are given in the Appendix.

## 2 Preliminaries

Let us consider a clinical experiment where patients come sequentially to the trial and are assigned to one of  $K \geq 2$  available treatments. The primary outcome  $Y_{kj}$  of the  $j$ th patient assigned to the  $k$ th treatment is usually assumed to be a continuous variable with

$$Y_{kj} = \mu_k + \epsilon_{kj}, \quad k = 1, \dots, K, \quad j = 1, \dots, N_k, \quad (1)$$

where  $\mu_k$  is the effect of the  $k$ th treatment,  $\epsilon_{kj}$ s are independent random errors with  $\epsilon_{kj} \sim N(0, \sigma^2)$  and  $\sigma^2$  denotes the unknown common variance. Letting  $N_k$  be the number of patients assigned to treatment  $k$  and  $n = \sum_{k=1}^K N_k$  the total sample size, then  $\pi_k = N_k/n$  is the proportion of patients assigned to the  $k$ th treatment; so let

$\boldsymbol{\pi} = (\pi_1, \dots, \pi_K)^\top$ , then  $\boldsymbol{\pi}^\top \mathbf{1}_K = 1$  for every  $n$ , where  $\mathbf{1}_K$  denotes the  $K$ -dim vector of ones.

Letting  $\hat{\mu}_{kn}$  ( $k = 1, \dots, K$ ) be the MLE of  $\mu_k$  after  $n$  steps, i.e., the corresponding sample mean, then  $\boldsymbol{\mu}$  and  $\hat{\boldsymbol{\mu}}_n$  are the  $K$ -dim vectors collecting the treatment effects and their estimates, respectively. As is well-known, conditional on the design the MLEs are strongly consistent and asymptotically normal, i.e., as  $n \rightarrow \infty$ ,  $\hat{\boldsymbol{\mu}}_n \rightarrow \boldsymbol{\mu}$  *a.s.* and  $\sqrt{n}(\hat{\boldsymbol{\mu}}_n - \boldsymbol{\mu}) \xrightarrow{d} \mathcal{N}(\mathbf{0}_K, \mathbf{M}^{-1})$ , where  $\mathbf{M} = \mathbf{M}(\boldsymbol{\pi}) = \sigma^{-2} \text{diag}(\boldsymbol{\pi})$  is the so-called normalized Fisher information matrix (conditional on the design) associated with  $\boldsymbol{\mu}$ . Under this scenario, a natural design question is how to allocate the  $K$  treatments to  $n$  subjects in order to optimize the inferential precision and several authors suggested optimal targets  $\boldsymbol{\rho} = (\rho_1, \dots, \rho_K)^\top$  (with  $\rho_k \geq 0$  and  $\boldsymbol{\rho}^\top \mathbf{1}_K = 1$ ) - as actual allocation proportions out of  $n$  patients or asymptotic limiting proportions for sequential trials - in order to optimize the estimation of the treatment effects. On the basis of the optimal design theory, these targets were derived by minimizing suitable convex criteria measuring the loss of information; classical examples are  $D$ -optimality and  $tr$ -optimality, intended to minimize the determinant  $|\cdot|$  or the trace  $\text{tr}(\cdot)$  of  $\mathbf{M}^{-1}(\boldsymbol{\rho})$ . By a well-known result (see, e.g., Silvey<sup>2</sup>), the balanced design  $\boldsymbol{\rho}^B = K^{-1} \mathbf{1}_K$  is both  $D$ - and  $tr$ -optimal.

In many circumstances, especially in the context of multiarm clinical trials, the inferential attention is devoted to the contrasts. So, letting  $\mathbf{A}^\top = [\mathbf{1}_{K-1} \mid -\mathbf{I}_{K-1}]$ , the  $(K-1)$ -dim vector of the contrasts with respect to the first treatment (considered as the reference) is  $\boldsymbol{\mu}_c = \mathbf{A}^\top \boldsymbol{\mu} = (\mu_1 - \mu_2, \dots, \mu_1 - \mu_K)^\top$ . The corresponding MLE is  $\hat{\boldsymbol{\mu}}_{cn} = \mathbf{A}^\top \hat{\boldsymbol{\mu}}_n$ , which is still strongly consistent and asymptotically normal with

$$\sqrt{n}(\hat{\boldsymbol{\mu}}_{cn} - \boldsymbol{\mu}_c) = \sqrt{n} \mathbf{A}^\top (\hat{\boldsymbol{\mu}}_n - \boldsymbol{\mu}) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{A}^\top \mathbf{M}^{-1} \mathbf{A}). \quad (2)$$

Within this framework, the balanced design  $\boldsymbol{\rho}^B$  is the so-called  $D_A$ -optimal allocation, which minimizes the determinant  $|\mathbf{A}^\top \mathbf{M}^{-1} \mathbf{A}|$ , while the  $tr_A$ -optimal design minimizing the trace of the asymptotic variance-covariance matrix in (2) is  $\boldsymbol{\rho}^R = (\rho_1^R, \dots, \rho_K^R)^\top$  given by<sup>13</sup>:

$$\rho_1^R = \frac{1}{1 + \sqrt{K-1}} \quad \text{and} \quad \rho_k^R = \frac{1}{\sqrt{K-1} (1 + \sqrt{K-1})}, \quad k = 2, \dots, K. \quad (3)$$

It is worth noticing that both  $\boldsymbol{\rho}^B$  and  $\boldsymbol{\rho}^R$  are motivated only by estimation precision without taking into account ethics.

On the other hand, Atkinson<sup>18</sup> proposed a target allocation intended to skew the assignments towards the best treatment in order to minimize the exposure of patients

to inefficacious treatments. In particular, assuming that a higher response is more desirable for patients and letting  $\bar{\mu} = K^{-1} \sum_{k=1}^K \mu_k$  be the overall treatments mean, target  $\rho^A$  proposed by Atkinson is

$$\rho_k^A = \Phi \left( \frac{\mu_k - \bar{\mu}}{\tau} \right) / \left[ \sum_{i=1}^K \Phi \left( \frac{\mu_i - \bar{\mu}}{\tau} \right) \right], \quad k = 1, \dots, K, \quad (4)$$

where  $\Phi(\cdot)$  is the cdf of a standard normal distribution and  $\tau > 0$  is a randomization parameter. Clearly, small values of  $\tau$  induce a strong ethical skew at the expense of estimation precision, while as  $\tau$  increases  $\rho^A$  tends to  $\rho^B$  and therefore more emphasis is given to inferential purposes.

### 3 Optimal target allocations for hypothesis testing

In what follows, we take into account the problem of testing hypothesis on the equality of the treatments effects by considering the null hypothesis of homogeneity  $H_0 : \mu_1 - \mu_2 = \dots = \mu_1 - \mu_K$ , i.e.,  $H_0 : \boldsymbol{\mu}_c = \mathbf{0}_{K-1}$ , versus the alternative  $H_A : \boldsymbol{\mu}_c \neq \mathbf{0}_{K-1}$ , where  $\mathbf{0}_{K-1}$  is the  $(K-1)$ -dim vector of zeros. **In addition to its intrinsic validity and generality, this overall homogeneity hypothesis represents the first stage of multiple comparison techniques of a wide class of step-down procedures applied in practice<sup>7;26</sup>, as will be discussed in Section 5.**

The classical Wald test statistic is

$$W_n = \frac{n}{\hat{\sigma}_n^2} \hat{\boldsymbol{\mu}}_{cn}^\top [\mathbf{A}^\top \text{diag}(\boldsymbol{\pi})^{-1} \mathbf{A}]^{-1} \hat{\boldsymbol{\mu}}_{cn},$$

where  $\hat{\sigma}_n^2$  is the usual consistent estimator of the common variance  $\sigma^2$ . Under the null hypothesis  $H_0$ ,  $W_n$  tends asymptotically to a chi-squared random variable with  $K-1$  degrees of freedom, while under the alternative hypothesis it converges to a non-central chi-squared r.v. with  $K-1$  degrees of freedom having non-centrality parameter

$$\phi(\boldsymbol{\pi}) = \frac{n}{\sigma^2} \boldsymbol{\mu}_c^\top [\mathbf{A}^\top \text{diag}(\boldsymbol{\pi})^{-1} \mathbf{A}]^{-1} \boldsymbol{\mu}_c.$$

The following Lemma provides a simple expression for the non-centrality parameter.



**Lemma 1.** *After  $n$  patients, the non-centrality parameter  $\phi(\boldsymbol{\pi})$  of Wald test could be rewritten as follows:*

$$\phi(\boldsymbol{\pi}) = \frac{n}{\sigma^2} \left\{ \sum_{k=1}^K (\mu_1 - \mu_k)^2 \pi_k - \left[ \sum_{k=1}^K (\mu_1 - \mu_k) \pi_k \right]^2 \right\}. \quad (5)$$

**Proof.** See Appendix 6.1.

From now on we adopt “the-larger-the-better” scenario, namely from an ethical viewpoint we assume that a higher response is more desirable. Moreover, for ease of notation and without loss of generality, we assume the following ordering between the treatment effects  $\mu_1 \geq \mu_2 \geq \dots \geq \mu_K$  (i.e., the best treatment will be labeled as the first one, while the  $K$ th treatment as the worst, admitting also clusters of treatments with the same efficacy). We stress that this choice is a simple label-coding, intended to avoid more complex notation; clearly, the treatment ranking is a-priori unknown but it can be estimated step-by-step in a sequential fashion, as discussed in Section 3.1.

**Remark 1.** *Letting  $\Delta_k = \mu_1 - \mu_k$  be the  $k$ th contrast wrt the best treatment (where, clearly,  $\Delta_1 = \mu_1 - \mu_1 = 0$ ), for a given target allocation  $\boldsymbol{\rho}$  the non-centrality parameter in (5) can be written as*

$$\phi(\boldsymbol{\rho}) = \frac{n}{\sigma^2} \left\{ \sum_{k=1}^K \Delta_k^2 \rho_k - \left[ \sum_{k=1}^K \Delta_k \rho_k \right]^2 \right\} = \frac{n}{\sigma^2} \sum_{1 \leq k < i \leq K} (\Delta_k - \Delta_i)^2 \rho_k \rho_i, \quad (6)$$

i.e., it is proportional to the variance of a discrete and non-negative random variable, say  $\Delta$ , with  $K$  (possibly different) ordered support points  $\Delta_K \geq \dots \geq \Delta_2 \geq \Delta_1 = 0$ , evaluated with respect to the pdf  $\boldsymbol{\rho}$ . Notice that, from the RHS of (6), the non-centrality parameter does not formally depend on the chosen reference treatment and it can be re-expressed in terms of every desirable treatment, say the  $r$ th, since  $\Delta_k - \Delta_i = \mu_i - \mu_k = (\mu_r - \mu_k) - (\mu_r - \mu_i)$ .

In the following Theorem we derive the optimal target allocation for  $K$  treatments that maximizes the power of the Wald test of homogeneity.

**Theorem 1.** *If  $\mu_1 > \mu_2 \geq \dots \geq \mu_{K-1} > \mu_K$ , then the optimal allocation maximizing the non-centrality parameter (6) of Wald test of homogeneity is  $\boldsymbol{\rho}^* = (1/2, 0, \dots, 0, 1/2)^\top$ , namely the target which takes into account only the best and the worst treatments in a balanced way. Moreover, in the case of a cluster*

of several best treatments  $\mu_1 = \mu_2 = \dots = \mu_j$  ( $j = 2, \dots, K - 1$ ), every allocation  $\rho^* = (\rho_1^*, \dots, \rho_K^*)^\top$  such that  $\rho_K^* = 1/2$  and  $\sum_{i=1}^j \rho_i^* = 1/2$  is optimal (i.e., any combination of the allocations of the treatments within the cluster with global probability  $1/2$  is optimal). The case of a cluster of several worst treatments can be derived analogously.

**Proof.** See Appendix 6.2.

Theorem 1 shows how an unconstrained maximization of the non-centrality parameter leads to a target  $\rho^*$  whose characteristics are inadequate both from the ethical and the inferential point of views. Indeed, even if  $\rho^*$  maximizes the power of Wald test,  $\rho^*$  does not allow to collect information on the efficacy of intermediate treatments, making the variance of the estimates unreliable. At the same time, problems also arise from the ethical viewpoint, since half of the patients are assigned to the less effective treatment. Furthermore, notice that every  $\rho_j^*$  is a discontinuous function of the unknown treatment effects which vanishes at the intermediate treatments and this does not allow the application of the classical asymptotic theory for response-adaptive procedures.

Taking into account two distinct objectives related to the ethical demands of skewing the allocations towards the superior treatment and the inferential goal of maximizing the inferential precision, several authors suggested multipurpose methodologies inspired by constrained/combined optimization approaches. In the context of binary trials, Tymofyeyev et al.<sup>7</sup> derived the allocation maximizing the power of Wald test of homogeneity subject to the ethical constraint that the total expected number of failures not exceed a given threshold. Since the corresponding solution involves only the best and the worst treatments with no assignments to the intermediate drugs, the authors superimposed an additional constraint by fixing a lower bound for each treatment allocation proportion. Although this approach is apparently flexible, the ensuing target is a discontinuous function of the unknown parameters, with a closed-form solution not generally available. Moreover, the subjective choice of these two thresholds could not be easy to interpret and could lead to formal contradictions, since they should depend in general on the unknown parameters<sup>16</sup>. The same problems may also occur in the combined optimization approach, where the choice of the weights assigned to ethics and inference plays a crucial role strongly affecting the results<sup>11</sup>.

To overcome the aforementioned drawbacks, in what follows we propose a new multi-objective design strategy based on a constrained optimization technique without superimposing any subjective choices. In particular, from the assumed order between the treatment effects  $\mu_1 \geq \mu_2 \geq \dots \geq \mu_K$ , in the next Theorem we derive

the constrained optimal allocation  $\tilde{\rho} = (\tilde{\rho}_1, \dots, \tilde{\rho}_K)^\top$  (with  $\tilde{\rho}_k \geq 0$  for every  $k = 1, \dots, K$  and  $\sum_{k=1}^K \tilde{\rho}_k = 1$ ) that maximizes the non-centrality parameter of the Wald test under the ethical constraint  $\tilde{\rho}_1 \geq \tilde{\rho}_2 \geq \dots \geq \tilde{\rho}_K$ .

**Theorem 2.** *If  $\mu_1 > \mu_2 \geq \dots \geq \mu_K$ , then the optimal constrained allocation  $\tilde{\rho} = (\tilde{\rho}_1, \dots, \tilde{\rho}_K)^\top$  maximizing the non-centrality parameter of the multivariate Wald test subject to the ethical constraint  $\tilde{\rho}_1 \geq \tilde{\rho}_2 \geq \dots \geq \tilde{\rho}_K$  is:*

$$\tilde{\rho} = \begin{cases} (1 - t[K - 1], t, \dots, t)^\top & \text{if } t \leq K^{-1}, \\ \rho^B & \text{if } t > K^{-1}, \end{cases}$$

where

$$t = \frac{\sum_{k=1}^K (\mu_1 - \mu_k)^2}{2 \left[ \sum_{k=1}^K (\mu_1 - \mu_k) \right]^2}. \quad (7)$$

*In the presence of a cluster of superior treatments  $\mu_1 = \dots = \mu_j > \mu_{j+1} \geq \dots \geq \mu_K$  (with  $j = 2, \dots, K - 1$ ), then  $\tilde{\rho} = \rho^B$  when  $t > K^{-1}$ , while for  $t \leq K^{-1}$  every allocation  $\tilde{\rho} = (\tilde{\rho}_1, \dots, \tilde{\rho}_j, t, \dots, t)^\top$  such that  $\tilde{\rho}_1 \geq \dots \geq \tilde{\rho}_j \geq t$  and  $\sum_{i=1}^j \tilde{\rho}_i = 1 - (K - j)t$  is optimal.*

**Proof.** See Appendix 6.3.

**Remark 2.** *In the same setting of Remark 1, the skewing parameter  $t$  in (7) could be rewritten as follows:*

$$t = \frac{\sum_{k=1}^K \Delta_k^2}{2 \left[ \sum_{k=1}^K \Delta_k \right]^2} = \frac{CV^2(\Delta) + 1}{2K},$$

where  $CV(\Delta)$  is the coefficient of variation of the random variable  $\Delta$  evaluated wrt the uniform distribution. Thus, on the basis of the values of the unknown treatment contrasts, the optimal constrained allocation  $\tilde{\rho}$  coincides with the balanced target if  $CV(\Delta) > 1$ ; whereas, when  $CV(\Delta) \leq 1$ , then  $t \leq K^{-1}$  and therefore  $\tilde{\rho}$  skews the assignments to the superior drug, while it maintains all the other treatments equireplicated.

From Theorem 2, the constrained optimal allocation has the following interesting properties:

- P1: through the skewing parameter  $t = t(\mu)$ ,  $\tilde{\rho}$  is a continuous and non-degenerate function of the unknown treatments effects and therefore it can be approached by standard response-adaptive randomization procedures;**

- P2: adopting  $\tilde{\rho}$ , the allocation proportion  $\tilde{\rho}_1$  to the best treatment is increasing as its effect  $\mu_1$  grows; more in general, every component  $\tilde{\rho}_k$  (for  $k = 2, \dots, K$ ) is a monotonically increasing function of  $\mu_k$ , provided that  $\mu_k \geq \bar{\mu}$  (see Appendix 6.4);
- P3: excluding the degenerate case  $\mu_1 = \dots = \mu_K$ , from (7) it can be shown that  $t \geq [2(K-1)]^{-1}$ , where the minimum value of the skewing parameter  $t$  is attained when  $\mu_1 > \mu_2 = \dots = \mu_K$  (see Case 3 of Appendix 6.3); in such a case the non-centrality parameter attains its maximum value, namely  $\phi(\tilde{\rho}) = \phi(\rho^*)$ ;
- P4:  $\tilde{\rho}$  is invariant under label permutation of the treatments, namely it does not change if the treatment labels are switched, provided that  $\mu_1$  is replaced by  $\max\{\mu_1, \dots, \mu_K\}$  in (7) (clearly, the target allocation proportion corresponding to the best treatment is  $1 - t[K-1]$ , while the other drugs are still equireplicated).

**Example 1.** To understand the behavior of the optimal constrained target  $\tilde{\rho}$  defined in Theorem 2, assume now  $K = 3$  treatments. From (7), it can be seen that  $t \leq 1/3$  if and only if  $(\mu_1 - \mu_2)(\sqrt{3} + 1) \geq (\mu_1 - \mu_3)(\sqrt{3} - 1)$ , namely when

$$\frac{\mu_1 - \mu_2}{\mu_1 - \mu_3} \geq \frac{\sqrt{3} - 1}{\sqrt{3} + 1} \simeq 0.268,$$

i.e., if  $\mu_1 - \mu_2 \geq 0.366 \cdot (\mu_2 - \mu_3)$ . Thus,  $\tilde{\rho}$  skews the assignments to the best treatment only when its relative superiority wrt the second drug - compared to the relative superiority between the other two treatments - is greater than the threshold 0.366, otherwise the allocations will be balanced. It is worth noticing that, even if  $t \leq 1/3$  induces a skewed allocation to the best treatment, at the same time from P3 the skewing parameter is always greater than  $[2(K-1)]^{-1} = 1/4$ ; therefore  $\tilde{\rho}_2 = \tilde{\rho}_3 \geq 1/4$  and  $\tilde{\rho}_1 \leq 1/2$ , which allows one to obtain valid performance also in terms of estimation precision. Table 1 summarizes the behavior of  $\tilde{\rho}$  as the treatment effects vary.

$\mu$			$\tilde{\rho}$			$t$
$\mu_1$	$\mu_2$	$\mu_3$	$\tilde{\rho}_1$	$\tilde{\rho}_2$	$\tilde{\rho}_3$	
12	10	1	1/3	1/3	1/3	0.370
12	9	1	0.336	0.332	0.332	0.332
12	6	1	0.457	0.272	0.272	0.272
12	3	1	0.495	0.253	0.253	0.253
12	1	1	0.500	0.250	0.250	0.250

**Table 1.** The behaviour of the optimal constrained target  $\tilde{\rho}$  with  $K = 3$  treatments as  $\mu_2$  varies.

### 3.1 Implementing optimal target allocations via RA procedures

The above-mentioned optimal constrained target depends on the unknown model parameters and therefore it is itself unknown at the beginning of the trial. So this allocation cannot be implemented directly, but it can be approached asymptotically by using suitable RA randomization procedures, namely sequential allocation rules that, in order to converge to the chosen target, change at each step the probabilities of treatment assignments on the basis of earlier responses and past allocations.

In particular, after a starting sample of  $n_0$  subjects assigned to each treatment - usually made via restricted randomization - to obtain non-trivial estimates of the unknown parameters, at each step  $n > Kn_0$  the treatment effects are estimated by  $\hat{\mu}_n = (\hat{\mu}_{1n}, \dots, \hat{\mu}_{Kn})^\top$  (where, according to our notation, the estimates are ranked from the largest - coded as the first one - to the smallest, namely the  $K$ th one, i.e.,  $\hat{\mu}_{1n} \geq \dots \geq \hat{\mu}_{Kn}$ ). Thus, the skewing parameter  $t$  in (7) is estimated by means of

$$\hat{t}_n = \frac{\sum_{k=1}^K (\hat{\mu}_{1n} - \hat{\mu}_{kn})^2}{2 \left[ \sum_{k=1}^K (\hat{\mu}_{1n} - \hat{\mu}_{kn}) \right]^2} \quad (8)$$

and the the optimal constrained target is estimated accordingly by

$$\hat{\rho}_n = \begin{cases} (1 - \hat{t}_n[K - 1], \hat{t}_n, \dots, \hat{t}_n)^\top & \text{if } \hat{t}_n \leq K^{-1}, \\ \rho^B & \text{if } \hat{t}_n > K^{-1}. \end{cases}$$

Thus, the next assignment is (randomly) forced to progressively approach the target. For instance, adopting the well-known sequential maximum likelihood design<sup>27</sup>, the  $(n + 1)$ th treatment is assigned with probabilities given by the current estimate of the target. While adopting the doubly-adaptive biased coin

design<sup>14</sup>, the treatment assignment is forced to the target increasingly as the distance between the current allocation proportion and the estimated target grows. As is well-known<sup>16</sup>, from P1 the treatment allocation proportion converges almost surely to  $\tilde{\rho}$  and the consistency of the estimators of the treatment effects is guaranteed along with their asymptotic normality. Thus, the asymptotic chi-squared approximation of Wald test follows directly; for a detailed discussion about the choice of  $n_0$  and the accuracy of the asymptotic approximation of Wald test statistic, see Baldi Antognini et al.<sup>28</sup>.

## 4 Comparison among target allocations

This section is dedicated to the performance comparison of different targets. As stressed by many authors, the interplay between different efficiency criteria (statistical power, efficient estimation and ethical considerations) is complex and does not lead to a unique best choice. To assess the performance with respect to different efficiency measures, we take into account several criteria, both from statistical and ethical point of views. More specifically, we consider the following design criteria:

- *Ethical demand*: as a measure of ethics we take into account the total expected responses, which is clearly maximized by assigning all subjects to the best treatment (i.e., in our setting, the first one); thus, given a target  $\rho$  we consider the ratio between the total expected outcomes and its optimal value  $\mu_1$ , i.e.,  $E_E(\rho) = \sum_{k=1}^K \mu_k \rho_k / \mu_1$ ;
- *Statistical power*: given the optimal target  $\rho^*$  maximizing the non-centrality parameter of Wald test (see Theorem 1), the measure of power-efficiency of an allocation  $\rho$  is evaluated by  $E_P(\rho) = \phi(\rho) / \phi(\rho^*)$ .

**Theorem 3.** *The constrained optimal target dominates the balanced design in terms of power and ethics, namely  $\tilde{\rho}$  guarantees higher values of  $E_P$  and  $E_E$  wrt  $\rho^B$ , simultaneously.*

**Proof.** See Appendix 6.5.

We compare the behavior of the constrained optimal target  $\tilde{\rho}$  and the unconstrained optimal one  $\rho^*$ , with three different allocations, namely the  $D_A$ -optimal target  $\rho^B$ , the  $tr_A$ -optimal allocation  $\rho^R$  in (3) and Atkinson's  $\rho^A$  in (4) with  $\tau = 1$  and  $\tau = 3$  (denoted by  $\rho_1^A$  and  $\rho_3^A$ , respectively). The aforementioned allocations have been proposed in order to meet specific objectives and therefore their behavior could vary with respect to different efficiency measures. More specifically,  $\rho^A$  has been suggested

for ethical reasons,  $\rho^R$  and  $\rho^B$  optimize a single design criterion for estimation goals,  $\rho^*$  maximizes the power of the Wald test without ethical restrictions and  $\tilde{\rho}$  is derived primarily for testing purposes but, at the same time, it tries to cope with the ethical issues raised from the unconstrained maximization.

To stress the impact of a given target in terms of estimation precision, in our comparisons we will also take into account the  $tr_A$ - and the  $D_A$ -efficiency, given respectively by

$$E_{tr_A}(\rho) = \frac{\text{tr}(\mathbf{A}^\top \mathbf{M}^{-1}(\rho^R) \mathbf{A})}{\text{tr}(\mathbf{A}^\top \mathbf{M}^{-1}(\rho) \mathbf{A})} \quad \text{and} \quad E_{D_A}(\rho) = \left\{ \frac{|\mathbf{A}^\top \mathbf{M}^{-1}(\rho^B) \mathbf{A}|}{|\mathbf{A}^\top \mathbf{M}^{-1}(\rho) \mathbf{A}|} \right\}^{\frac{1}{K-1}}.$$

The following tables summarize the results for  $K = 3$  and  $K = 5$  treatments, varying the values of the treatment effects. Note that both  $\rho^R$  and  $\rho^B$  depend only on the number  $K$  of considered drugs and therefore they do not change as  $\mu$  varies, while  $\rho^*$  always assigns 50% of the allocations both to the best and the worst treatment. It is worth noting that, although these targets do not depend on  $\mu$ , the corresponding efficiency measures  $E_E$  and  $E_P$  change as the treatment effects vary.

Insert Table 2 and Table 3 here

Let us first consider the results in Table 2 with  $K = 3$  treatments. As far as the statistical power is concerned,  $\tilde{\rho}$  is the closest to the corresponding optimum target  $\rho^*$ , followed by  $\rho^R$ . The gain in terms of power of  $\tilde{\rho}$  wrt  $\rho^R$  ranges from about 1% to about 4%. Adopting the allocation  $\rho^A$ , the choice of the randomization parameter  $\tau$  is crucial: for  $\tau = 1$  the ethical skew is strong, so that the assignment to the worst treatment vanishes in almost every parameter setting; while  $\tau = 3$  gives more emphasis to inference. In general,  $\rho^A$  shows the lowest statistical power but, at the same time, it exhibits the highest ethical efficiency, followed by  $\tilde{\rho}$  (except for the scenario  $\mu = (12, 11, 1)^\top$  under which  $\tilde{\rho} = \rho^B$ ). In general, the ethical impact of  $\tilde{\rho}$  is superior than that of  $\rho^R$ , with a gain up to 7%. For what concerns the  $tr_A$ -efficiency and the  $D_A$ -efficiency,  $\tilde{\rho}$  is always close to the optimum, with maximum losses lower than 3% and 8%, respectively. As stated previously,  $\rho^*$  and  $\rho_1^A$  are strongly inadequate from the viewpoint of estimation precision.

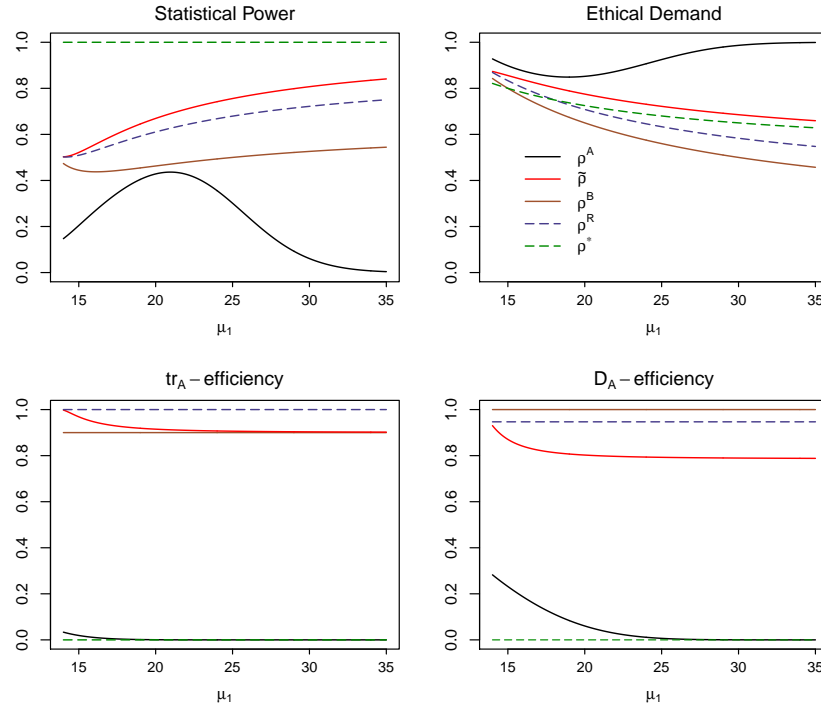
Table 3 refers to the scenario with  $K = 5$  treatments and the results substantially reflect those obtained for  $K = 3$ . More specifically, in terms of power,  $\tilde{\rho}$  is still the closest to the optimum  $\rho^*$ ; if compared to  $\rho^R$ , the third best option,  $\tilde{\rho}$  induces a gain of efficiency up to 6%. As for  $K = 3$ ,  $\rho^A$  tends to have the best performance in terms of ethics, except for some cases where  $\tilde{\rho}$  has substantially the same ethical efficiency

wrt  $\rho_3^A$  (notice that the choice of  $\tau = 1$  does not always guarantee a superior ethical efficiency wrt  $\tau = 3$ ). Whereas,  $\tilde{\rho}$  guarantees an ethical gain up to about 8% wrt  $\rho^R$ .

Recalling that our proposal represents a valid compromise between ethics and power, in the case of  $K = 5$  treatments,  $\tilde{\rho}$  becomes less efficient wrt  $\rho^R$  in terms of estimation precision; however, the constrained optimal target still ensures very good performances, especially in terms of  $tr_A$ -optimality with an efficiency always greater than 91%, while the  $D_A$ -efficiency is always greater than 80%. Indeed, following the results of Baldi Antognini and Giovagnoli<sup>11</sup>, it is easy to show that  $\tilde{\rho}$  is  $D_A$ -admissible, namely it does not exist another target which is superior wrt both ethics and  $D_A$ -optimality, simultaneously. Once again,  $\rho^*$  and  $\rho_1^A$  exhibit values of  $tr_A$ -efficiency and  $D_A$ -efficiency close to zero.

In order to explore the behavior of the considered targets also from a graphical point of view, Figure 1 shows their performance in the case of  $K = 5$  treatments with  $\mu^\top = (\mu_1, 13, 12, 11, 9)$  as  $\mu_1$  varies between 14 and 35. Essentially, this Figure confirms the above-mentioned results, also showing an anomalous behavior of Atkinson's target, which exhibits a non-monotone efficiency in terms of both statistical power and ethical demands as the effect of the best treatment grows. Indeed, as  $\mu_1$  increases,  $\rho_1^A$  tends to assign all subjects to the best treatment; contrary to the other targets under which the allocation proportion to the worst drug does not vanish. This explains why the power of  $\rho_1^A$  tends to vanish, while its ethical efficiency tends to one. For what concerns the remaining targets, they have a decreasing ethical efficiency due to the denominator  $\mu_1$  of the considered ethical measure. In general,  $\tilde{\rho}$  has better performance wrt all of them. Also in this setting, adopting  $\rho^*$  and  $\rho_1^A$  the estimation precision vanishes.





**Figure 1.** Efficiency measures for  $\rho_1^A$ ,  $\tilde{\rho}$ ,  $\rho^B$ ,  $\rho^R$  and  $\rho^*$  with  $\mu^\top = (\mu_1, 13, 12, 11, 9)$  as  $\mu_1$  varies between 14 and 35.

Finally, note that  $\rho^R$  assigns more subjects to the reference treatment of the contrasts, which in our assumptions also coincides with the superior treatment; thus, the ethical efficiency of  $\rho^R$  is emphasized under the previous scenarios. Whereas, if  $\mu^\top = (1, 3, 6)$  (i.e., the best treatment is the third one), then  $\rho^R = (0.414, 0.293, 0.293)^\top$  with  $E_E(\rho^R) = 0.508$ , while  $\tilde{\rho} = (0.266, 0.266, 0.468)^\top$  has ethical efficiency  $E_E(\tilde{\rho}) = 0.646$ .

As stressed by Sverdlov and Rosenberger<sup>13</sup>, a target showing high efficiency under one criterion may have low efficiency under other criteria, since statistical and ethical purposes can directly compete with one another, resulting in a complex pattern. However, the optimal constrained target  $\tilde{\rho}$  shows good performance both from statical and ethical point of views, since the efficiency gains obtained in statistical power are not achieved at the expense of poor ethical demands, guaranteing at the same time valid performance in terms of estimation precision.

## 5 Discussion

The present paper deals with the problem of finding suitable target allocations for multiarm clinical trials with normally responses. To achieve an appropriate trade-off among inferential precision and ethical concerns, we take into account the inferential goal of maximizing the power of the test of homogeneity between the treatment effects, under a suitable ethical constraint reflecting the effectiveness of the treatments. The ensuing optimal constrained target, which guarantees very good performances in terms of both ethical gain and statistical efficiency, is a non-degenerate continuous function of the unknown contrasts that can be approached by standard RA randomization procedures.

Besides its generality, the considered overall null hypothesis of homogeneity represents a cornerstone in the statistical literature and constitutes the first stage of multiple comparison methodologies for several stepwise procedures. For instance, Fisher's least significant difference method is a two-step test for pairwise comparisons which is considered one of the most powerful methodologies for comparing several treatments. In the first step, the overall null hypothesis of homogeneity is tested and, in the case of rejection, the second step is performed in order to test the pairwise comparisons at the same level of significance<sup>29</sup>. Since in some practical applications not all the comparisons may be of primary interest and the inferential goal consists in testing the superiority of a set of new drugs with respect to the placebo (i.e., by considering a right-tailed alternative hypothesis), Zhu and Wong<sup>30</sup> derived the optimal compound target for normal homoscedastic outcomes by combining into a single design criterion the variances of the usual  $t$ -tests for pairwise comparisons with different weights reflecting the relative importance of each of them. Clearly, this approach does not take into account ethical concerns and, in the case of equal weights (namely when all the  $K - 1$  comparisons have the same importance), this compound optimal target coincides with  $\rho^R$  in (3).

From a design perspective, the normal linear model is a standard framework by which approximating new phenomena. Such a model tends to be appropriate also for real non-normal data by applying suitable transformations to the responses (e.g., by the well-known power transformation<sup>23</sup>). However, in order to extend the applicability of the suggested methodology, one of the future directions of our work is to relax the assumption of homoscedasticity, taking also into account covariates/prognostic factors that potentially affect the outcomes. In such cases,

---

**the level of complexity of the problem will increase significantly and we hope to find out a tractable solution in the future.**

## Acknowledgements

Research supported by the Italian Ministry of Education, University and Research under PRIN 2015 “Environmental processes and human activities: capturing their interactions via statistical methods (EphaStat)”. The authors are grateful to the referees for their helpful comments, which led to a substantially improved version of the paper. Moreover, we wish to thank Andrea Ghiglietti and Giacomo Aletti for their suggestions.

## References

1. Efron B. Forcing sequential experiments to be balanced. *Biometrika* 1971; 58: 403–417.
2. Silvey SD. *Optimal Designs*. Chapman & Hall, London, 1980.
3. Begg CB and Kalish LA. Treatment allocation for nonlinear models in clinical trials: The logistic model. *Biometrics* 1984; 40: 409–420.
4. Kalish LA and Harrington DP. Efficiency of balanced treatment allocation for survival analysis. *Biometrics* 1988; 44: 409–420.
5. Baldi Antognini A. A theoretical analysis of the power of biased coin designs. *Journal of Statistical Planning and Inference* 2008; 138: 1792–1798.
6. Azriel D, Mandel M and Rinott Y. Optimal allocation to maximize power of two-sample tests for binary response. *Biometrika* 2012; 99: 101–113.
7. Tymofyeyev Y, Rosenberger WF and Hu F. Implementing optimal allocation in sequential binary response experiments. *Journal of the American Statistical Association* 2007; 102: 224–234.
8. Rosenberger WF, Stallard N, Ivanova A et al. Optimal adaptive designs for binary response trials. *Biometrics* 2001; 57: 909–913.
9. Biswas A, Bhattacharya R and Zhang L. Optimal response-adaptive designs for continuous responses in phase III trials. *Biometrical Journal* 2007; 49: 928–940.
10. Biswas A and Bhattacharya R. Optimal response-adaptive designs for normal responses. *Biometrical Journal* 2009; 51: 193–202.
11. Baldi Antognini A and Giovagnoli A. Compound optimal allocation for individual and collective ethics in binary clinical trials. *Biometrika* 2010; 97: 935–946.

12. Baldi Antognini A and Zagoraiou M. Multi-objective optimal designs in comparative clinical trials with covariates: the reinforced doubly-adaptive biased coin design. *The Annals of Statistics* 2012; 40: 1315–1345.
13. Sverdlov O and Rosenberger WF. On recent advances in optimal allocation designs in clinical trials. *Journal of Statistical Theory and Practice* 2013; 7(4): 753–773.
14. Hu F and Zhang LX. Asymptotic properties of doubly adaptive biased coin designs for multi-treatment clinical trials. *The Annals of Statistics* 2004; 32: 268–301.
15. Atkinson AC and Biswas A. *Randomised Response-Adaptive Designs in Clinical Trials*. Boca Raton: Chapman & Hall/CRC Press, 2014.
16. Baldi Antognini A and Giovagnoli A. *Adaptive Designs for Sequential Treatment Allocation*. Chapman & Hall/CRC Biostatistics, 2015.
17. Rosenberger WF and Lachin JL. *Randomization in Clinical Trials: Theory and Practice, Second Edition*. John Wiley & Sons, New York, 2015.
18. Atkinson AC. Adaptive biased-coin designs for clinical trials with several treatments. *Discussiones Mathematicae Probability and Statistics* 2004; 24(1): 85–108.
19. Atkinson AC and Biswas A. Bayesian adaptive biased-coin designs for clinical trials with normal responses. *Biometrics* 2005; 61: 118–125.
20. Sverdlov O, Tymofyeyev Y and Wong WK. Optimal response-adaptive randomized designs for multi-armed survival trials. *Statistics in Medicine* 2011; 30(24): 2890–2910.
21. Wilson TW, Lacourciere Y and Barnes CC. The antihypertensive efficacy of losartan and amlodipine assessed with office and ambulatory blood pressure monitoring. *Canadian Medical Association Journal* 1998; 159: 469–476.
22. Prasad AS, Fitzgerald JT, Bao B et al. Duration of symptoms and plasma cytokine levels in patients with the common cold treated with zinc acetate. *Annals of internal medicine* 2000; 133: 245–252.
23. Atkinson AC and Biswas A. Adaptive biased-coin designs for skewing the allocation proportion in clinical trials with normal responses. *Statistics in Medicine* 2005; 24: 2477–2492.
24. Dworkin RH, Corbin AE, Young JP et al. Pregabalin for the treatment of postherpetic neuralgia. a randomized, placebo-controlled trial. *Neurology* 2003; 60: 1274–1283.
25. Zhang LX and Rosenberger WF. Response-adaptive randomization for clinical trials with continuous outcomes. *Biometrics* 2006; 62: 562–569.
26. Christensen R. *Plain answers to complex questions: the theory of linear models*. Springer-Verlag, New York, 2002.
27. Melfi V and Page C. Estimation after adaptive allocation. *Journal of Statistical Planning and Inference* 2000; 29: 353–363.

28. Baldi Antognini A, Vagheggini A and Zagoraiou M. Is the classical wald test always suitable under response-adaptive randomization? *Statistical Methods in Medical Research* 2016; available online: 1–18.
29. Hochberg Y and Tamhane AC. *Multiple Comparison Procedures*. Wiley, New York, 1987.
30. Zhu W and Wong WK. Optimal treatment allocation in comparative biomedical studies. *Statistics in Medicine* 2000; 19: 639–648.

## 6 Appendix

Given a discrete r.v.  $\Delta$  with values  $\Delta_1, \dots, \Delta_K$ , from now on we set  $M_\rho(\Delta^s) = \sum_{k=1}^K \Delta_k^s \rho_k$  for  $s = 1, 2$ , so that  $M_\rho(\Delta) = \sum_{k=1}^K \Delta_k \rho_k$  and  $V_\rho(\Delta) = M_\rho(\Delta^2) - [M_\rho(\Delta)]^2$  are the expectation and the variance of  $\Delta$  evaluated wrt the pdf  $\rho$ . To avoid a cumbersome notation, the same quantities evaluated wrt  $\rho^B$  will be denoted by  $M(\cdot)$  and  $V(\cdot)$ .

### 6.1 Proof of Lemma 1

Note that

$$\mathbf{A}^\top \text{diag}(\boldsymbol{\pi})^{-1} \mathbf{A} = \text{diag}(\pi_k^{-1})_{k=2, \dots, K} + \pi_1^{-1} \mathbf{1}_{K-1} \mathbf{1}_{K-1}^\top = \mathbf{T} + \mathbf{u} \mathbf{v}^\top,$$

where  $\mathbf{T} = \text{diag}(\pi_k^{-1})_{k=2, \dots, K}$ ,  $\mathbf{u} = \mathbf{1}_{K-1}$  and  $\mathbf{v} = \pi_1^{-1} \mathbf{1}_{K-1}$ . Moreover,  $\mathbf{T} + \mathbf{u} \mathbf{v}^\top$  is invertible since  $1 + \mathbf{v}^\top \mathbf{T}^{-1} \mathbf{u} = 1 + \pi_1^{-1} \mathbf{1}_{K-1}^\top \mathbf{T}^{-1} \mathbf{1}_{K-1} = \pi_1^{-1} \neq 0$ ; therefore, by applying the Sherman-Morrison formula,

$$(\mathbf{T} + \mathbf{u} \mathbf{v}^\top)^{-1} = \mathbf{T}^{-1} - \frac{\mathbf{T}^{-1} \mathbf{u} \mathbf{v}^\top \mathbf{T}^{-1}}{1 + \mathbf{v}^\top \mathbf{T}^{-1} \mathbf{u}}.$$

Moreover,  $\boldsymbol{\mu}_c^\top \mathbf{T}^{-1} \boldsymbol{\mu}_c = \sum_{k=2}^K \Delta_k^2 \pi_k$ ,  $\boldsymbol{\mu}_c^\top \mathbf{T}^{-1} \mathbf{u} = \sum_{k=2}^K \Delta_k \pi_k$  and  $\mathbf{v}^\top \mathbf{T}^{-1} \boldsymbol{\mu}_c = \pi_1^{-1} \sum_{k=2}^K \Delta_k \pi_k$ ; thus,

$$\boldsymbol{\mu}_c^\top \left[ \frac{\mathbf{T}^{-1} \mathbf{u} \mathbf{v}^\top \mathbf{T}^{-1}}{1 + \mathbf{v}^\top \mathbf{T}^{-1} \mathbf{u}} \right] \boldsymbol{\mu}_c = \left( \sum_{k=2}^K \Delta_k \pi_k \right)^2$$

and therefore the non-centrality parameter in (5) follows via simple algebra.

### 6.2 Proof of Theorem 1

As shown in Remark 1, the non-centrality parameter of Wald test of homogeneity could be rewritten as  $\phi(\boldsymbol{\rho}) = n\sigma^{-2}V_\rho(\Delta)$ , where  $\Delta$  is a non-negative discrete r.v.

with  $K$  (possibly different) ordered support points  $0 = \Delta_1 \leq \Delta_2 \leq \dots \leq \Delta_K$ . By a well-known result, for a given support the variance is maximized by the pdf that assigns half mass of probability at the two extremes; moreover, this maximum is global and is unique. Since  $\mu_1 > \mu_2$  and  $\mu_{K-1} > \mu_K$ , then  $\Delta_K > \Delta_{K-1} \geq \dots \geq \Delta_2 > \Delta_1 = 0$  and therefore the non-centrality parameter  $\phi$  is maximized by  $\boldsymbol{\rho}^* = (1/2, 0, \dots, 0, 1/2)^\top$ . If there exists a cluster of treatments such that  $\mu_1 = \mu_2 = \dots = \mu_j$  (with  $j < K$ ), then the first  $j$  support points collapse into a single value since  $0 = \Delta_1 = \dots = \Delta_j$  and therefore every combination  $(\rho_1^*, \dots, \rho_j^*)$  with  $\rho_k^* \geq 0$  and  $\sum_{k=1}^j \rho_k^* = 1/2$  is optimal; clearly, the same reasoning still holds for more than one worst treatments.

### 6.3 Proof of Theorem 2

We now want to derive the target allocation  $\tilde{\boldsymbol{\rho}} = (\tilde{\rho}_1, \dots, \tilde{\rho}_K)^\top$  maximizing the non-centrality parameter of Wald test of homogeneity under the ethical constraints  $\tilde{\rho}_1 \geq \tilde{\rho}_2 \geq \dots \geq \tilde{\rho}_K$ . As shown in Remark 1,  $\phi(\boldsymbol{\rho}) = n\sigma^{-2}V_\rho(\Delta)$  and, since  $n\sigma^{-2}$  is a constant, the problem is finding the constrained target  $\tilde{\boldsymbol{\rho}}$  that maximizes  $V_\rho(\Delta) = M_\rho(\Delta^2) - [M_\rho(\Delta)]^2 = \sum_{1 \leq k < i \leq K} (\Delta_k - \Delta_i)^2 \rho_k \rho_i$ . The corresponding Lagrange function is

$$\mathcal{L}(\boldsymbol{\rho}, \lambda_1, \dots, \lambda_K) = \sum_{1 \leq k < i \leq K} (\Delta_k - \Delta_i)^2 \rho_k \rho_i - \sum_{j=1}^{K-1} \lambda_j (\rho_{j+1} - \rho_j) - \lambda_K \left( \sum_{j=1}^K \rho_j - 1 \right).$$

If we let now  $a_i = \sum_{k=1}^K (\Delta_k - \Delta_i)^2 \rho_k$  for  $i = 1, \dots, K$ , by putting the partial derivatives of the Lagrangian equal to zero, we obtain the following system of equations:

$$\begin{cases} a_1 + \lambda_1 = \lambda_K \\ a_2 + \lambda_2 - \lambda_1 = \lambda_K \\ \vdots \\ a_{K-1} + \lambda_{K-1} - \lambda_{K-2} = \lambda_K \\ a_K - \lambda_{K-1} = \lambda_K. \end{cases}$$

By summing up all the equations,  $(a_1 + \lambda_1) + (a_2 + \lambda_2 - \lambda_1) + \dots + (a_{K-1} + \lambda_{K-1} - \lambda_{K-2}) + (a_K - \lambda_{K-1}) = K\lambda_K$ , which gives  $\lambda_K = K^{-1} \sum_{k=1}^K a_k = \bar{a} >$

0, so that the previous system could be rewritten as follows

$$\begin{cases} \lambda_1 = \bar{a} - a_1 \\ \lambda_2 = \lambda_1 + \lambda_K - a_2 = (\bar{a} - a_1) + (\bar{a} - a_2) \\ \vdots \\ \lambda_{K-1} = \sum_{i=1}^{K-1} (\bar{a} - a_i) \\ \lambda_K = \bar{a}, \end{cases}$$

where  $a_1 = \sum_{k=1}^K \Delta_k^2 \rho_k = M_\rho(\Delta^2)$ ,

$$a_k = a_1 + \Delta_k^2 - 2\Delta_k M_\rho(\Delta) = V_\rho(\Delta) + [\Delta_k - M_\rho(\Delta)]^2, \quad k = 1, \dots, K \quad (9)$$

and

$$\bar{a} = a_1 + M(\Delta^2) - 2M(\Delta)M_\rho(\Delta) = V_\rho(\Delta) + V(\Delta) + [M_\rho(\Delta) - M(\Delta)]^2. \quad (10)$$

Since  $\mu_1 > \mu_K$ , then  $0 = \Delta_1 < \Delta_K$  and therefore  $V(\Delta) > 0$ , which implies that  $\bar{a} > 0$ , namely  $\lambda_K > 0$ .

**Case 1:**  $\lambda_i > 0$  for every  $i = 1, \dots, K - 1$ .

Conditions  $\lambda_i > 0 \forall i$  imply that  $\rho_1 = \dots = \rho_K = 1/K$ , namely the corresponding target is the balanced one  $\rho^B$ . Under  $\rho^B$ , from (9) and (10) it follows that  $a_k = V(\Delta) + [\Delta_k - M(\Delta)]^2$  for every  $k = 1, \dots, K$  and  $\bar{a} = 2V(\Delta)$ . Condition  $\lambda_1 > 0$ , namely  $\bar{a} > a_1$ , implies that

$$2V(\Delta) > M(\Delta^2) \Leftrightarrow V(\Delta) > M(\Delta)^2 \Leftrightarrow CV(\Delta) > 1;$$

while conditions  $\lambda_i > 0$  for every  $i = 2, \dots, K - 1$  correspond to  $\sum_{k=1}^i (\bar{a} - a_k) > 0$ , namely  $i\bar{a} > \sum_{k=1}^i a_k$ . Moreover, from (9),  $a_i \geq a_{i+1} \Leftrightarrow 2M(\Delta)[\Delta_{i+1} - \Delta_i] \geq \Delta_{i+1}^2 - \Delta_i^2$ , i.e.,  $(\Delta_i + \Delta_{i+1})/2 \leq M(\Delta)$  with  $\Delta_{i+1} > 0$ . Since  $0 = \Delta_1 \leq \Delta_2 \leq \dots \leq \Delta_K$ , then  $\exists \tilde{i}$  such that  $a_1 \geq a_2 \geq \dots \geq a_{\tilde{i}} \leq a_{\tilde{i}+1} \leq \dots \leq a_K$ , namely the sequence  $\{a_k; k = 1, \dots, K\}$  is decreasing for any  $k \leq \tilde{i}$ , while it is increasing for every  $k > \tilde{i}$ . Therefore, if  $CV(\Delta) > 1$  then  $\bar{a} > a_1 \geq a_2 \geq \dots \geq a_{\tilde{i}}$  which guarantees that  $\lambda_i > 0$  for  $i = 1, \dots, \tilde{i}$ . Whereas, for  $i > \tilde{i}$  the sub-sequence  $\{i^{-1} \sum_{k=1}^i a_k; i = \tilde{i} + 1, \dots, K - 1\}$  becomes increasing and therefore if

$$\bar{a} > \frac{\sum_{k=1}^{K-1} a_k}{K-1} \quad (11)$$

then  $\lambda_i > 0$  for every  $i = \tilde{i} + 1, \dots, K - 1$ . Condition (11) could be rewritten as follows

$$2V(\Delta) > \frac{\sum_{k=1}^{K-1} \{V(\Delta) + [\Delta_k - M(\Delta)]^2\}}{K - 1}$$

namely

$$(K - 1)V(\Delta) > \sum_{k=1}^{K-1} [\Delta_k - M(\Delta)]^2 = KV(\Delta) - [\Delta_K - M(\Delta)]^2,$$

which is clearly satisfied if

$$V(\Delta) < [\Delta_K - M(\Delta)]^2. \quad (12)$$

Recalling that  $CV(\Delta) > 1$ , from Popoviciu's inequality it follows that  $M(\Delta)^2 < V(\Delta) \leq \Delta_K^2/4$ , which implies that  $\Delta_K > 2M(\Delta)$ ; therefore (12) is satisfied, since in this case  $\Delta_K^2/4 < [\Delta_K - M(\Delta)]^2$ .

**Case 2:**  $\lambda_1 = 0$  and  $\lambda_i > 0$  for every  $i = 2, \dots, K - 1$ .

Since  $\lambda_i > 0$  for every  $i = 2, \dots, K$ , then  $\rho_2 = \rho_3 = \dots = \rho_K$ , namely the corresponding target should assume the form

$$\boldsymbol{\rho}^\top = (1 - (K - 1)t, t, \dots, t, t) \quad \text{with} \quad 0 \leq t \leq \frac{1}{K}. \quad (13)$$

Condition  $\lambda_1 = 0$ , i.e.,  $\bar{a} = a_1$ , implies that

$$V_\rho(\Delta) + V(\Delta) + [M_\rho(\Delta) - M(\Delta)]^2 = M_\rho(\Delta^2) \Leftrightarrow 2M_\rho(\Delta)M(\Delta) = M(\Delta^2); \quad (14)$$

however, under (13),  $M_\rho(\Delta) = \sum_{k=2}^K \Delta_k t = tKM(\Delta)$  and therefore, from (14),  $2Kt[M(\Delta)]^2 = M(\Delta^2)$ , namely

$$t = \frac{M(\Delta^2)}{2K[M(\Delta)]^2}, \quad (15)$$

which is an admissible solution only when  $t \leq K^{-1}$ , i.e., for  $CV(\Delta) \leq 1$ . Given this solution, we have now to check if conditions  $\lambda_i > 0$  for every  $i = 2, \dots, K - 1$  hold true. These conditions are clearly satisfied when  $\mu_1 > \mu_2$  because in this case, from



$\bar{a} = a_1$  and (9),

$$\begin{aligned} a_1 > a_2 &\Leftrightarrow 2M_\rho(\Delta)\Delta_2 > \Delta_2^2 \\ 2a_1 > a_2 + a_3 &\Leftrightarrow 2M_\rho(\Delta)[\Delta_2 + \Delta_3] > \Delta_2^2 + \Delta_3^2 \\ &\vdots \\ (K-2)a_1 &> \sum_{k=2}^{K-1} a_k \Leftrightarrow 2M_\rho(\Delta) \sum_{k=2}^{K-1} \Delta_k > \sum_{k=2}^{K-1} \Delta_k^2, \end{aligned}$$

where  $M_\rho(\Delta) = tKM(\Delta) = M(\Delta^2)/[2M(\Delta)]$ . Therefore, these conditions could be rewritten as follows

$$\frac{M(\Delta^2)}{M(\Delta)} > \frac{\Delta_2^2}{\Delta_2}, \quad \frac{M(\Delta^2)}{M(\Delta)} > \frac{\Delta_2^2 + \Delta_3^2}{\Delta_2 + \Delta_3}, \dots, \frac{M(\Delta^2)}{M(\Delta)} > \frac{\sum_{k=2}^{K-1} \Delta_k^2}{\sum_{k=2}^{K-1} \Delta_k}, \quad (16)$$

which are clearly satisfied provided that  $0 < \Delta_2 < \Delta_K$ . Whereas,

- if  $\mu_1 = \mu_2$  (or, more in general, in the presence of a cluster of better treatments), then  $\Delta_2 = 0$  and therefore  $\bar{a} = a_1 = a_2$ , namely  $\lambda_2 = 0$ , contradicting the assumptions;
- if  $\mu_1 > \mu_2 = \dots = \mu_K$ , namely  $\Delta_2 = \dots = \Delta_K = \alpha > 0$ , then  $M(\Delta) = \alpha(1 - 1/K)$  and  $M(\Delta^2) = \alpha^2(1 - 1/K)$ , so that  $M(\Delta^2)/M(\Delta) = \alpha$  which contradicts (16), since we should obtain  $M(\Delta^2)/M(\Delta) = \alpha > \alpha$ .

**Case 3:**  $\lambda_1 = \dots = \lambda_{K-1} = 0$ .

Under this scenario,  $\bar{a} = a_1$ , so  $2M_\rho(\Delta) = M(\Delta^2)/M(\Delta)$  and  $a_1 = \dots = a_{K-1}$ , namely  $2M_\rho(\Delta) = \Delta_2 = \dots = \Delta_{K-1}$ . By combining the previous conditions,  $\Delta_2 = \dots = \Delta_{K-1} = \Delta_K = \alpha > 0$ , so that  $M(\Delta) = \alpha(1 - 1/K)$  and  $M(\Delta^2) = \alpha^2(1 - 1/K)$  and thus  $2M_\rho(\Delta) = 2 \sum_{k=1}^K \Delta_k \rho_k = \alpha$ , i.e.,  $2\alpha \sum_{k=2}^K \rho_k = \alpha$ . Therefore,  $\rho_1 = 1/2$  and the ensuing optimal target is  $\rho^\top = (2^{-1}; [2(K-1)]^{-1}; \dots; [2(K-1)]^{-1})$ . Notice that this scenario represents a special case of the optimal target derived in Case 2; indeed, when  $\mu_1 > \mu_2 = \dots = \mu_K$ , then  $CV(\Delta) = (K-1)^{-1/2}$  and  $t = [2(K-1)]^{-1}$ , which is the minimum value for the skewing parameter, since

$$t = \frac{\sum_{k=2}^K \Delta_k^2}{2 \left[ \sum_{k=2}^K \Delta_k \right]^2} \geq \frac{1}{2(K-1)} \Leftrightarrow (K-1) \sum_{k=2}^K \Delta_k^2 \geq \left[ \sum_{k=2}^K \Delta_k \right]^2, \quad (17)$$

where the RHS trivially holds from Jensen's inequality.

**Case 4:**  $\lambda_1 > 0$  and at least one multiplier  $\lambda_i = 0$ ,  $2 \leq i \leq K - 1$ .

If  $\lambda_i = 0$  with  $2 \leq i \leq K - 1$ , then  $i\bar{a} = a_1 + \dots + a_i$ . From (9) and (10),

$$ia_1 + i[M(\Delta^2) - 2M(\Delta)M_\rho(\Delta)] = ia_1 + \Delta_2^2 + \dots + \Delta_i^2 - 2M_\rho(\Delta)(\Delta_2 + \dots + \Delta_i),$$

namely

$$2M_\rho(\Delta) = \frac{iM(\Delta^2) - (\Delta_2^2 + \dots + \Delta_i^2)}{iM(\Delta) - (\Delta_2 + \dots + \Delta_i)}. \quad (18)$$

Since  $\lambda_1 > 0$ , then  $\bar{a} > a_1$  and therefore  $a_2 + \dots + a_i > (i-1)\bar{a} > (i-1)a_1$ .

However, note that

$$a_2 + \dots + a_i > (i-1)a_1 \Leftrightarrow 2M_\rho(\Delta) < \frac{\Delta_2^2 + \dots + \Delta_i^2}{\Delta_2 + \dots + \Delta_i} \quad (19)$$

where  $\sum_{k=2}^i \Delta_k > 0$ , namely at least  $\Delta_i > 0$ . Thus, combining (18) and (19),

$$\frac{iM(\Delta^2) - (\Delta_2^2 + \dots + \Delta_i^2)}{iM(\Delta) - (\Delta_2 + \dots + \Delta_i)} < \frac{\Delta_2^2 + \dots + \Delta_i^2}{\Delta_2 + \dots + \Delta_i},$$

namely

$$(\Delta_2 + \dots + \Delta_i) \sum_{k=2}^K \Delta_k^2 < (\Delta_2^2 + \dots + \Delta_i^2) \sum_{k=2}^K \Delta_k,$$

i.e.

$$(\Delta_2 + \dots + \Delta_i) \sum_{k=i+1}^K \Delta_k^2 < (\Delta_2^2 + \dots + \Delta_i^2) \sum_{k=i+1}^K \Delta_k,$$

which could be rewritten as follows

$$\sum_{k=i+1}^K \Delta_k \left[ \sum_{l=2}^i \Delta_l (\Delta_k - \Delta_l) \right] < 0$$

and it is impossible since  $\Delta_k \geq \Delta_l \geq 0$  for any  $k > l$ .

**Case 5:**  $\lambda_1 = 0$  with at least one  $\lambda_i > 0$  and  $\lambda_{i+1} = 0$  for  $i \in [2; K - 2]$ .

Under this setting  $\bar{a} = a_1$  and  $i\bar{a} > a_1 + \dots + a_i$ , namely

$$2M_\rho(\Delta) = \frac{M(\Delta^2)}{M(\Delta)} > \frac{\sum_{k=2}^i \Delta_k^2}{\sum_{k=2}^i \Delta_k},$$

with  $\Delta_i > 0$ . However, at the same time  $(i+1)\bar{a} = a_1 + \dots + a_{i+1}$ , so that

$$2M_\rho(\Delta) = \frac{\sum_{k=2}^{i+1} \Delta_k^2}{\sum_{k=2}^{i+1} \Delta_k} = \frac{M(\Delta^2)}{M(\Delta)},$$

which implies that  $\Delta_2 = \dots = \Delta_K = \alpha > 0$ . However, this scenario is impossible since

$$2M_\rho(\Delta) = \alpha > \frac{\sum_{k=2}^i \Delta_k^2}{\sum_{k=2}^i \Delta_k} = \alpha.$$

**Case 6:**  $\lambda_1 = 0$  with  $\lambda_2 = \dots = \lambda_j = 0$  and  $\lambda_i > 0$  for  $i = j+1, \dots, K-1$ .

Under this setting  $\bar{a} = a_1$ ,  $j\bar{a} = a_1 + \dots + a_j$  and  $(j+1)\bar{a} > a_1 + \dots + a_j + a_{j+1}$ , namely  $\bar{a} = a_1 > a_{j+1}$  and therefore  $2M_\rho(\Delta) > \Delta_{j+1} > 0$ . At the same time,  $j\bar{a} = ja_1 = \sum_{k=1}^j a_k \Leftrightarrow (j-1)a_1 = \sum_{k=2}^j a_k$ , so that:

- if  $\sum_{k=2}^j \Delta_k > 0$ , namely at least  $\Delta_j > 0$ , then

$$2M_\rho(\Delta) = \frac{\sum_{k=2}^j \Delta_k^2}{\sum_{k=2}^j \Delta_k} > \Delta_{j+1},$$

which is clearly impossible, since  $\Delta_{j+1} \geq \Delta_k$  for every  $k = 2, \dots, j$  and therefore  $\sum_{k=2}^j \Delta_k^2 < \Delta_{j+1} \sum_{k=2}^j \Delta_k$ .

- if  $\sum_{k=2}^j \Delta_k = 0$  then  $\Delta_2 = \dots = \Delta_j = 0$  and, since  $\lambda_i > 0$  for  $i = j+1, \dots, K-1$ , the corresponding target has the form  $\rho^\top = (x_1, \dots, x_j, t, \dots, t)$ , with  $0 \leq t \leq K^{-1}$ ,  $x_1 \geq \dots \geq x_j \geq t$  and  $\sum_{k=1}^j x_k = 1 - (K-j)t$ . Therefore, under this setting  $2M_\rho(\Delta) = 2t \sum_{k=j+1}^K \Delta_k = 2tKM(\Delta)$  and, since  $\bar{a} = a_1$ , then  $2M_\rho(\Delta) = M(\Delta^2)/M(\Delta)$ , so that  $t = M(\Delta^2)/(2K[M(\Delta)]^2)$  is an admissible solution only when  $t \leq K^{-1}$ , namely if  $CV(\Delta) \leq 1$ . For the same reasons of Case 2, it can be easily shown that the conditions  $\lambda_i > 0$  for  $i = j+1, \dots, K-1$  are trivially satisfied.

## 6.4 Proof of Property P2

Clearly if  $CV(\Delta) > 1$ , i.e. when the optimal constrained target coincides with the balanced one, then all the partial derivatives vanish. Whereas, assuming now that  $CV(\Delta) \leq 1$ , then  $\tilde{\rho}_1$  is increasing in  $\mu_1$ ; indeed,

$$\frac{\partial \tilde{\rho}_1}{\partial \mu_1} = -(K-1) \left[ \frac{\left[ \sum_{k=2}^K (\mu_1 - \mu_k) \right]^2 - (K-1) \sum_{k=2}^K (\mu_1 - \mu_k)^2}{\left[ \sum_{k=2}^K (\mu_1 - \mu_k) \right]^3} \right]$$

is non-negative, since from (17),  $(K-1) \sum_{k=2}^K \Delta_k^2 \geq \left[ \sum_{k=2}^K \Delta_k \right]^2$ . Taking into account  $\tilde{\rho}_j$  ( $j = 2, \dots, K$ ), then

$$\frac{\partial \tilde{\rho}_j}{\partial \mu_j} = \frac{\sum_{i=1}^K (\mu_1 - \mu_i)^2 - (\mu_1 - \mu_j) \sum_{i=1}^K (\mu_1 - \mu_i)}{\left[ \sum_{i=1}^K (\mu_1 - \mu_i) \right]^3}$$

is non-negative if and only if  $M(\Delta^2) \geq M(\Delta)\Delta_j$ , namely when  $V(\Delta) \geq M(\Delta)[\Delta_j - M(\Delta)]$ . Clearly, the previous condition is verified for every treatment  $j$  such that  $\Delta_j \leq M(\Delta)$ , namely when  $\mu_j \geq \bar{\mu}$ .

### 6.5 Proof of of Theorem 3

Adopting the balanced target, then  $E_E(\rho^B) = \bar{\mu}/\mu_1$  and

$$E_P(\rho^B) = \frac{V(\Delta)}{\left(\frac{\Delta_K}{2}\right)^2}.$$

Taking into account  $\tilde{\rho}$ , if  $t > K^{-1}$  then  $\tilde{\rho} = \rho^B$ ; whereas, when  $t \leq K^{-1}$  the efficiency measure of ethics is  $E_E(\tilde{\rho}) = 1 - tK + tK(\bar{\mu}/\mu_1)$  and the power-efficiency becomes

$$E_P(\tilde{\rho}) = \frac{V_{\tilde{\rho}}(\Delta)}{\left(\frac{\Delta_K}{2}\right)^2} = \frac{\left(\frac{M(\Delta^2)}{2M(\Delta)}\right)^2}{\left(\frac{\Delta_K}{2}\right)^2},$$

since, from (6) and (15),

$$V_{\tilde{\rho}}(\Delta) = tK \{M(\Delta^2) - tK[M(\Delta)]^2\} = \left(\frac{M(\Delta^2)}{2M(\Delta)}\right)^2.$$

Therefore  $E_E(\tilde{\rho}) \geq E_E(\rho^B)$ , because  $\mu_1(1 - tK) \geq \bar{\mu}(1 - tK)$ , since  $tK \leq 1$  and  $\mu_1 \geq \bar{\mu}$ . Moreover,  $E_P(\tilde{\rho}) \geq E_P(\rho^B)$ , due to the fact that

$$\left(\frac{M(\Delta^2)}{2M(\Delta)}\right)^2 = \left\{ \frac{V(\Delta) + [M(\Delta)]^2}{2M(\Delta)} \right\}^2 \geq V(\Delta),$$

namely  $\{V(\Delta) + [M(\Delta)]^2\}^2 \geq 4[M(\Delta)]^2V(\Delta)$ , since  $\{V(\Delta) - [M(\Delta)]^2\}^2 \geq 0$ .

**Table 2.** The case of  $K = 3$  treatments:  $\rho^B = (1/3, 1/3, 1/3)^\top$ ,  $\rho^* = (0.5, 0, 0.5)^\top$  and  $\rho^R = (0.414, 0.293, 0.293)^\top$ .

$\mu^\top$	Targets	$E_E(\rho)$	$E_F(\rho)$	$E_{tr_A}(\rho)$	$E_{D_A}(\rho)$
(6, 3, 1)	$\rho_1^A = (0.724, 0.269, 0.007)^\top$	0.860	0.302	0.080	0.194
	$\rho_3^A = (0.547, 0.306, 0.147)^\top$	0.724	0.591	0.849	0.815
	$\tilde{\rho} = (0.468, 0.266, 0.266)^\top$	0.646	0.722	0.988	0.945
	$\rho^B$	0.556	0.676	0.971	1
	$\rho^R$	0.609	0.715	1	0.979
	$\rho^*$	0.643	1	$\rightarrow 0$	$\rightarrow 0$
(9, 3, 1)	$\rho_1^A = (0.916, 0.083, 0.001)^\top$	0.944	0.174	0.005	0.029
	$\rho_3^A = (0.671, 0.234, 0.095)^\top$	0.759	0.614	0.656	0.635
	$\tilde{\rho} = (0.49, 0.255, 0.255)^\top$	0.603	0.797	0.978	0.928
	$\rho^B$	0.481	0.722	0.971	1
	$\rho^R$	0.544	0.780	1	0.979
	$\rho^*$	0.556	1	$\rightarrow 0$	$\rightarrow 0$
(12, 3, 1)	$\rho_1^A = (0.99, 0.01, 0)^\top$	0.993	0.026	$\rightarrow 0$	0.001
	$\rho_3^A = (0.771, 0.171, 0.058)^\top$	0.819	0.533	0.454	0.454
	$\tilde{\rho} = (0.494, 0.253, 0.253)^\top$	0.579	0.843	0.975	0.923
	$\rho^B$	0.444	0.757	0.971	1
	$\rho^R$	0.512	0.821	1	0.979
	$\rho^*$	0.542	1	$\rightarrow 0$	$\rightarrow 0$
(12, 1, 1)	$\rho_1^A = (1, 0, 0)^\top$	1	$\rightarrow 0$	$\rightarrow 0$	$\rightarrow 0$
	$\rho_3^A = (0.818, 0.091, 0.091)^\top$	0.833	0.597	0.478	0.429
	$\tilde{\rho} = (0.5, 0.25, 0.25)^\top$	0.542	1	0.971	0.919
	$\rho^B$	0.389	0.889	0.971	1
	$\rho^R$	0.463	0.971	1	0.979
	$\rho^*$	0.542	1	$\rightarrow 0$	$\rightarrow 0$
(12, 7, 1)	$\rho_1^A = (0.613, 0.387, 0)^\top$	0.839	0.196	$\rightarrow 0$	$\rightarrow 0$
	$\rho_3^A = (0.626, 0.354, 0.019)^\top$	0.835	0.240	0.200	0.339
	$\tilde{\rho} = (0.43, 0.285, 0.285)^\top$	0.620	0.688	0.999	0.971
	$\rho^B$	0.556	0.669	0.971	1
	$\rho^R$	0.609	0.688	1	0.979
	$\rho^*$	0.542	1	$\rightarrow 0$	$\rightarrow 0$
(12, 11, 1)	$\rho_1^A = (0.5, 0.5, 0)^\top$	0.958	0.008	$\rightarrow 0$	$\rightarrow 0$
	$\rho_3^A = (0.516, 0.478, 0.006)^\top$	0.955	0.028	0.063	0.193
	$\tilde{\rho} = \rho^B$	0.667	0.815	0.971	1
	$\rho^R$	0.707	0.773	1	0.979
	$\rho^*$	0.542	1	$\rightarrow 0$	$\rightarrow 0$
	(12,7,2)	$\rho_1^A = (0.667, 0.333, 0)^\top$	0.861	0.222	$\rightarrow 0$
$\rho_3^A = (0.635, 0.333, 0.032)^\top$		0.835	0.303	0.311	0.427
$\tilde{\rho} = (0.444, 0.278, 0.278)^\top$		0.653	0.694	0.996	0.962
$\rho^B$		0.583	0.667	0.971	1
$\rho^R$		0.634	0.692	1	0.979
$\rho^*$		0.583	1	$\rightarrow 0$	$\rightarrow 0$
(12,7,4)	$\rho_1^A = (0.798, 0.202, 0)^\top$	0.916	0.252	0.001	0.021
	$\rho_3^A = (0.639, 0.284, 0.076)^\top$	0.83	0.492	0.591	0.613
	$\tilde{\rho} = (0.473, 0.263, 0.263)^\top$	0.715	0.732	0.986	0.941
	$\rho^B$	0.639	0.681	0.971	1
	$\rho^R$	0.683	0.723	1	0.979
	$\rho^*$	0.667	1	$\rightarrow 0$	$\rightarrow 0$
(12,7,6)	$\rho_1^A = (0.908, 0.083, 0.009)^\top$	0.961	0.241	0.092	0.135
	$\rho_3^A = (0.619, 0.229, 0.152)^\top$	0.829	0.774	0.822	0.763
	$\tilde{\rho} = (0.496, 0.252, 0.252)^\top$	0.769	0.854	0.974	0.922
	$\rho^B$	0.694	0.765	0.971	1
	$\rho^R$	0.732	0.832	1	0.979
	$\rho^*$	0.750	1	$\rightarrow 0$	$\rightarrow 0$

Prepared using sagej.cls

**Table 3.** The case of  $K = 5$  treatments:  $\rho^B = (0.2, 0.2, 0.2, 0.2, 0.2)^\top$ ,  $\rho^* = (0.5, 0, 0, 0, 0.5)^\top$  and  $\rho^R = (1/3, 1/6, 1/6, 1/6, 1/6)^\top$ .

$\mu^\top$	Targets	$E_E(\rho)$	$E_P(\rho)$	$E_{tr_A}(\rho)$	$E_{D_A}(\rho)$
(14,13,12,11,9)	$\rho_1^A = (0.37, 0.332, 0.217, 0.08, 0.001)^\top$	0.928	0.147	0.034	0.282
	$\rho_3^A = (0.305, 0.26, 0.209, 0.157, 0.07)^\top$	0.893	0.321	0.847	0.867
	$\tilde{\rho} = (0.355, 0.161, 0.161, 0.161, 0.161)^\top$	0.873	0.503	0.998	0.930
	$\rho^B$	0.843	0.474	0.900	1
	$\rho^R$	0.869	0.502	1	0.947
	$\rho^*$	0.821	1	$\rightarrow 0$	$\rightarrow 0$
(16,13,12,11,9)	$\rho_1^A = (0.43, 0.339, 0.181, 0.05, 0)^\top$	0.876	0.264	0.011	0.186
	$\rho_3^A = (0.364, 0.246, 0.192, 0.14, 0.058)^\top$	0.837	0.392	0.807	0.813
	$\tilde{\rho} = (0.452, 0.137, 0.137, 0.137, 0.137)^\top$	0.837	0.554	0.947	0.840
	$\rho^B$	0.763	0.438	0.900	1
	$\rho^R$	0.802	0.528	1	0.947
	$\rho^*$	0.781	1	$\rightarrow 0$	$\rightarrow 0$
(18,13,12,11,9)	$\rho_1^A = (0.504, 0.33, 0.138, 0.028, 0)^\top$	0.851	0.367	0.003	0.112
	$\rho_3^A = (0.41, 0.235, 0.179, 0.126, 0.049)^\top$	0.801	0.479	0.751	0.760
	$\tilde{\rho} = (0.476, 0.131, 0.131, 0.131, 0.131)^\top$	0.803	0.618	0.925	0.814
	$\rho^B$	0.700	0.446	0.900	1
	$\rho^R$	0.750	0.572	1	0.947
	$\rho^*$	0.750	1	$\rightarrow 0$	$\rightarrow 0$
(20,13,12,11,9)	$\rho_1^A = (0.595, 0.297, 0.094, 0.014, 0)^\top$	0.852	0.428	0.001	0.060
	$\rho_3^A = (0.449, 0.227, 0.168, 0.115, 0.041)^\top$	0.779	0.550	0.690	0.710
	$\tilde{\rho} = (0.486, 0.129, 0.129, 0.129, 0.129)^\top$	0.775	0.669	0.915	0.803
	$\rho^B$	0.650	0.463	0.900	1
	$\rho^R$	0.708	0.611	1	0.947
	$\rho^*$	0.725	1	$\rightarrow 0$	$\rightarrow 0$
(18,13,12,11,2)	$\rho_1^A = (0.315, 0.304, 0.248, 0.133, 0)^\top$	0.781	0.118	$\rightarrow 0$	$\rightarrow 0$
	$\rho_3^A = (0.354, 0.26, 0.217, 0.169, 0)^\top$	0.789	0.130	0.014	0.253
	$\tilde{\rho} = (0.367, 0.158, 0.158, 0.158, 0.158)^\top$	0.701	0.453	0.995	0.921
	$\rho^B$	0.622	0.421	0.900	1
	$\rho^R$	0.685	0.451	1	0.947
	$\rho^*$	0.556	1	$\rightarrow 0$	$\rightarrow 0$
(18,13,12,11,4)	$\rho_1^A = (0.351, 0.323, 0.23, 0.096, 0)^\top$	0.796	0.155	$\rightarrow 0$	$\rightarrow 0$
	$\rho_3^A = (0.372, 0.257, 0.209, 0.159, 0.002)^\top$	0.795	0.175	0.073	0.382
	$\tilde{\rho} = (0.402, 0.149, 0.149, 0.149, 0.149)^\top$	0.734	0.467	0.981	0.890
	$\rho^B$	0.644	0.413	0.900	1
	$\rho^R$	0.704	0.460	1	0.947
	$\rho^*$	0.611	1	$\rightarrow 0$	$\rightarrow 0$
(18,13,12,11,6)	$\rho_1^A = (0.4, 0.337, 0.2, 0.063, 0)^\top$	0.815	0.213	$\rightarrow 0$	0.007
	$\rho_3^A = (0.391, 0.252, 0.2, 0.148, 0.009)^\top$	0.800	0.252	0.265	0.537
	$\tilde{\rho} = (0.436, 0.141, 0.141, 0.141, 0.141)^\top$	0.765	0.498	0.959	0.857
	$\rho^B$	0.667	0.411	0.900	1
	$\rho^R$	0.722	0.481	1	0.947
	$\rho^*$	0.667	1	$\rightarrow 0$	$\rightarrow 0$
(18,13,12,11,11)	$\rho_1^A = (0.587, 0.293, 0.093, 0.013, 0.013)^\top$	0.877	0.580	0.211	0.307
	$\rho_3^A = (0.409, 0.215, 0.159, 0.109, 0.109)^\top$	0.803	0.747	0.919	0.847
	$\tilde{\rho} = (0.492, 0.127, 0.127, 0.127, 0.127)^\top$	0.823	0.826	0.909	0.800
	$\rho^B$	0.722	0.555	0.900	1
	$\rho^R$	0.769	0.746	1	0.947
	$\rho^*$	0.806	1	$\rightarrow 0$	$\rightarrow 0$