Book of Short Papers SIS 2020

Editors: Alessio Pollice, Nicola Salvati and Francesco Schirripa Spagnolo

Published by Pearson

www.pearson.com

ISBN 9788891910776

Optimal designs for multi-arm exponential trials

Disegni ottimi per prove cliniche a risposte esponenziali

Rosamarie Frieri and Marco Novelli

Abstract Most of the randomized clinical trials for treatment comparisons have been designed to obtain a balanced allocation among the treatments. This is mostly due to the so-called universal optimality of balance. However, with several treatments the balanced allocation may not be efficient and could be strongly ethically inappropriate, in particular for phase III-trials. In [3], taking into account the exponential model, the target allocation maximizing the power of Wald test under a suitable ethical constraint has been derived. In this paper, we further explore the operating characteristics of such allocation through a comparison with other targets proposed in the literature, showing that the constrained optimal target exhibits good performances in terms of inferential precision and ethical demands.

Abstract *La maggior parte degli studi clinici randomizzati per il confronto tra trattamenti sono stati disegnati per ottenere un'allocazione bilanciata tra i gruppi. Questo è dovuto principalmente alla cosiddetta ottimalità universale del bilanciamento. Tuttavia, in presenza di molti trattamenti, l'allocazione bilanciata potrebbe risultare inefficiente e fortemente non etica, in particolare nelle prove cliniche di Fase III. In [3], considerando risposte esponenziali, è stata derivata l'allocazione ottimale che massimizza la potenza del test di Wald basato sui contrasti, sotto un appropriato vincolo etico. In questo articolo, verranno approfondite le caratteristiche operative di tale allocazione anche attraverso un confronto con le altre allocazioni proposte in letteratura, allo scopo di valutare la sua efficienza rispetto a criteri di natura sia etica che inferenziale.*

Key words: unbalanced allocations, efficiency, power of the Wald test, ethics

---

Rosamarie Frieri
Department of Statistical Sciences, University of Bologna, Via Belle Arti 41, Bologna
e-mail: rosamarie.frieri2@unibo.it

Marco Novelli
Department of Statistical Sciences, University of Bologna, Via Belle Arti 41, Bologna
e-mail: marco.novelli4@unibo.it

1 Introduction

In randomized clinical trials two competing goals, i.e. individual vs. collective ethics, have to be balanced. Indeed, the need to allocate more patients to the best available treatment (individual ethics) usually conflicts with the rigourous pursuit of scientific knowledge obtained with high inferential precision (collective ethics). So far, the vast majority of randomized clinical trials have been designed to achieve a balanced allocation among treatment groups, thanks to the optimal properties of balance especially in terms of estimation accuracy (see [6]). However, adopting equal allocation in the presence of several treatments could be neither efficient - since it is different from the optimal design for hypothesis testing (see [2, 3]) - nor ethical, especially in the context of phase-III trials, where the need to care for the well-being of the subjects involved in the trial is of primary importance. To overcome this trade-off, target allocations depending on a metric that accounts for treatment effects and/or their variabilities have been proposed, in order to obtain a valid compromise between ethical demand and inferential precision (see [1, 5]). Generally, these allocations depend on the unknown model parameters and can be targeted by suitable response adaptive randomization procedures, namely sequential allocation rules that, making use of the information accrued along the trial, change the assignment probabilities at each step in order to skew allocations toward the superior treatment. Recently, taking into account the problem of testing statistical hypothesis in normal homoscedastic trials, Baldi Antognini et. al. in [2] proposed an optimal target which maximizes the power of the Wald test of homogeneity, subject to an ethical constraint reflecting the effectiveness of the treatments. Moreover, Frieri and Zagoraiou in [3] derived a constrained target for exponential outcomes that are particularly relevant for oncological trials with survival endpoints. In this paper, we explore in depth the operating characteristic of the allocation derived in [3] through a comparison with other targets proposed in the literature. Our results show that the constrained optimal target guarantees very good performance in terms of statistical power, estimation precision, and ethical demands.

2 Framework and notation

In this work, clinical trials in which each subject is sequentially allocated to one of $K \geq 2$ available treatments are considered. Let δ_{kj} be the treatment assignment indicator such that $\delta_{kj} = 1$ when patients j is assigned to treatment $k\,(k = 1, \ldots, K)$ and 0 otherwise, with $\sum_{k=1}^{K} \delta_{kj} = 1$. The experimental outcome of the corresponding patient, Y_j, is assumed to be exponentially distributed with $E(Y_j|\delta_{kj} = 1) = \mu_k$, the treatment effect, and $V(Y_j|\delta_{kj} = 1) = \mu_k^2$ its variance. At each stage n, let $\boldsymbol{\pi}_n = (\pi_{1n}, \ldots, \pi_{Kn})^\top$, with $\pi_{kn} = n^{-1} \sum_{j=1}^{n} \delta_{kj}$ and $\sum_{k=1}^{K} \pi_{kn} = 1$, be the vector collecting the treatment assignment proportions up to that point, and $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_K)^\top$ the vectors of treatment effects. In what follows, without loss of generality, we adopt

the-larger-the-better scenario, that is an higher response is more desirable for the patient's care, and we will work under the non-restrictive assumption that $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_K$, i.e. the best performing treatment is labelled as the first one and the worst as the Kth one.

Here, the inferential focus is on the treatment contrasts, so letting $\mathbf{A}^\top = [\mathbf{1}_{K-1}| -\mathbf{I}_{K-1}]$, where $\mathbf{1}_p$ and $\mathbf{I}_p$ represent the p-dim vector of ones and the identity matrix, we denote by $\boldsymbol{\mu}_c = \mathbf{A}^\top \boldsymbol{\mu} = (\mu_1 - \mu_2, \ldots, \mu_1 - \mu_K)^\top$ and $\hat{\boldsymbol{\mu}}_{cn} = (\hat{\mu}_{1n} - \hat{\mu}_{2n}, \ldots, \hat{\mu}_{1n} - \hat{\mu}_{Kn})^\top$ the vector of contrasts wrt the first treatment and their MLEs, respectively. Under well-known regularity conditions, $\hat{\boldsymbol{\mu}}_{cn}$ is strongly consistent and asymptotically normal, i.e. $\hat{\boldsymbol{\mu}}_{cn} \xrightarrow{a.s.} \boldsymbol{\mu}_c$ and $\sqrt{n}(\hat{\boldsymbol{\mu}}_{cn} - \boldsymbol{\mu}_c) \xrightarrow{d} \mathrm{N}(\mathbf{0}_{K-1}, \mathbf{A}^\top \mathbf{M}^{-1} \mathbf{A})$, where $\mathbf{M} = \mathbf{M}(\boldsymbol{\mu}|\boldsymbol{\pi}) = \mathrm{diag}\left(\mu_i^{-2}\pi_i\right)_{i=1,\ldots,K}$ is the Fisher information matrix associated with $\boldsymbol{\mu}$. Finally, let us define by $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_K)^\top$, with $\rho_k \geq 0$ and $\sum_{k=1}^K \rho_k = 1$, the desired target allocation proportion, that can be obtained through suitable optimization problems.

The experimental strategy adopted to obtain the optimal target depends on the objective of the trial. When the aim is to maximize the inferential precision in the estimation of the treatment contrasts, Sverdlov and Rosenberger in [7] derived the tr_A optimal target by minimizing $\mathrm{tr}[\mathbf{A}^\top \mathbf{M}^{-1} \mathbf{A}]$, i.e.

$$\rho_1^A = \frac{\mu_1 \sqrt{K-1}}{\mu_1 \sqrt{K-1} + \sum_{i=2}^K \mu_i} \quad \text{and} \quad \rho_k^A = \frac{\mu_k}{\mu_1 \sqrt{K-1} + \sum_{i=2}^K \mu_i} \text{ for } k = 2, \ldots, K. \quad (1)$$

It is easy to show that $\rho_i^A \geq \rho_{i+1}^A \iff \mu_i \geq \mu_{i+1}$ for $i = 1, \ldots, K-1$ so this target, for exponential outcomes and under the-larger-the-better scenario, is ethical. On the side of hypothesis testing instead, a typical problem in multi-arm trials is to test the null-hypothesis of equality of treatment effects, i.e. $\boldsymbol{\mu}_c = \mathbf{0}_{K-1}$, where $\mathbf{0}_{K-1}$ is the $(K-1)$-dimensional vector of zeros. The target allocation maximizing the power of the Wald test of homogeneity is $\boldsymbol{\rho}^* = (\mu_1/(\mu_1+\mu_K), 0, \ldots, 0, \mu_K/(\mu_1+\mu_K))^\top$ (see [3]), which is clearly inappropriate for both statistical and ethical reasons. To avoid empty treatment arms, Zhu and Hu in [9], adopting the same framework in [8], set an optimization problem in which the power of the test is maximized subject to a constraint on the lower bound of the minimum number of subject assigned to each treatment. More specifically, the ensuing target $\boldsymbol{\rho}^Z$ should satisfy $\rho_i^Z \geq T$ for $i = 1, \ldots, K$, where $T \in [0, 1/K]$ is selected by the user. The target $\boldsymbol{\rho}^Z$ is available in closed form (which is not reported here for brevity, see [9]), however, it is only defined when $\mu_1 = \cdots = \mu_s > \mu_{s+1} \geq \ldots \geq \mu_{K-g} > \mu_{K-g+1} = \cdots = \mu_K$, for some positive integers s and g such that $s + g < K$. Notice that this framework does not include the configurations of the parameters in which $s + g = K$, i.e. $\mu_1 = \cdots = \mu_j > \mu_{j+1} = \cdots = \mu_K$ for $j = 2, \ldots, K-1$.

Finally, Frieri and Zagoraiou in [3], by adopting a multipurpose design methodology, derived the optimal allocation maximizing the power of Wald test subject to an ethical constraint reflecting the efficacy of the treatments. The ensuing constrained optimal target maximizing the non centrality parameter $\phi(\boldsymbol{\rho}) = n \cdot \mu_c^\top [\mathbf{A}^\top \mathbf{M}^{-1} \mathbf{A}]^{-1} \mu_c$ of the multivariate Wald test subject to $\rho_i \geq \rho_{i+1}$ for $i =$

$1,\dots,K-1$ is

$$\boldsymbol{\rho}^C = \begin{cases} (1-(K-1)x, x, \dots, x)^\top & \text{if } x < K^{-1}, \\ \boldsymbol{\rho}^B & \text{if } x \geq K^{-1}, \end{cases} \tag{2}$$

where $\boldsymbol{\rho}^B$ is the balanced allocation and

$$x = \frac{\frac{1}{\mu_1}\sum_{k=1}^{K}\left(\frac{1}{\mu_k}-\frac{1}{\mu_1}\right)^2}{\sum_{k=1}^{K}\left(\frac{1}{\mu_k}-\frac{1}{\mu_1}\right)\sum_{k=1}^{K}\left(\frac{1}{\mu_k^2}-\frac{1}{\mu_1^2}\right)}.$$

Note that in the presence of a cluster of superior treatments $\mu_1 = \dots = \mu_s > \mu_{s+1} \geq \dots \geq \mu_{K-g} > \mu_{K-g+1} = \dots = \mu_K$ all targets $\boldsymbol{\rho}^C = (\rho_1^C, \dots, \rho_s^C, x, \dots, x)^\top$ such that $\sum_{i=1}^{s}\rho_i^C = 1-(K-s)x$ and $s+g<K$ are optimal. Instead, when $s+g=K$ every allocation such that $\sum_{i=1}^{s}\rho_i^C = 1-(K-s)x = \mu_1/(\mu_1+\mu_K) = 1-\sum_{i=s+1}^{K}\rho_i^C$ is optimal.

3 Comparisons of optimal allocations for the exponential model

In this section, we compare the statistical performances of the previously introduced designs, that is $\boldsymbol{\rho}^A, \boldsymbol{\rho}^B, \boldsymbol{\rho}^C$ and $\boldsymbol{\rho}^Z$, in terms of the normalized power, $E_P(\boldsymbol{\rho}) = \phi(\boldsymbol{\rho})/\phi(\boldsymbol{\rho}^*)$ and the tr_A efficiency, $E_{tr_A}(\boldsymbol{\rho}) = \frac{\mathrm{tr}[\mathbf{A}^\top\mathbf{M}^{-1}(\boldsymbol{\rho}^A)\mathbf{A}]}{\mathrm{tr}[\mathbf{A}^\top\mathbf{M}^{-1}(\boldsymbol{\rho})\mathbf{A}]}$. As a measure of ethics, we consider the ratio between the total expected outcome and its maximum value, that is $E_E(\boldsymbol{\rho}) = \mu_1^{-1}\sum_{i=1}^{K}\mu_i\rho_i$. Figure 1 summarizes the operating characteristics of the targets for $\mu_2 = 10, \mu_3 = 9$ and $\mu_4 = 8$, as μ_1 varies from 15 to 35. As far as the statistical power is concerned, for values of μ_1 close to μ_2, $\boldsymbol{\rho}^Z$ with $T = 0.1$ shows the highest power efficiency while, as μ_1 increases (greater than 25) the best performance in terms of $E_P(\cdot)$ is achieved by $\boldsymbol{\rho}^C$ in (2), whose power is always increasing wrt μ_1. Note that, this property is not shared by all the targets. The $\boldsymbol{\rho}^A$ target in (1), the balanced one and $\boldsymbol{\rho}^Z$ for $T = 0.15$ and $T = 0.2$ always present lower power than $\boldsymbol{\rho}^C$. In terms of ethical demand, $\boldsymbol{\rho}^C$ outperforms all the competitors with a gain wrt the second best ($\boldsymbol{\rho}^Z$, $T = 0.1$) up to 7%. The ethical efficiency of $\boldsymbol{\rho}^C$ - as also confirmed by other studies omitted here for brevity - is slightly decreasing for values of μ_1 close to μ_2 and tends to increase as μ_1 grows. A similar behaviour is retrieved only for $\boldsymbol{\rho}^A$, whereas all the remaining targets have decreasing ethical efficiency as μ_1 increases. In terms of estimation precision, the second best is $\boldsymbol{\rho}^Z$ with $T = 0.15$ for $\mu_1 < 26$, while the same target with $T = 0.1$ shows a value of E_{tr_A} approaching 1 as μ_1 grows. The constrained target $\boldsymbol{\rho}^C$ exhibits an estimation efficiency almost constant wrt μ_1 with a value always greater than 0.93. In general, the balanced design shows the lowest efficiency in all the measures considered, while the performances of $\boldsymbol{\rho}^Z$ strongly depends on the subjective choice of T.

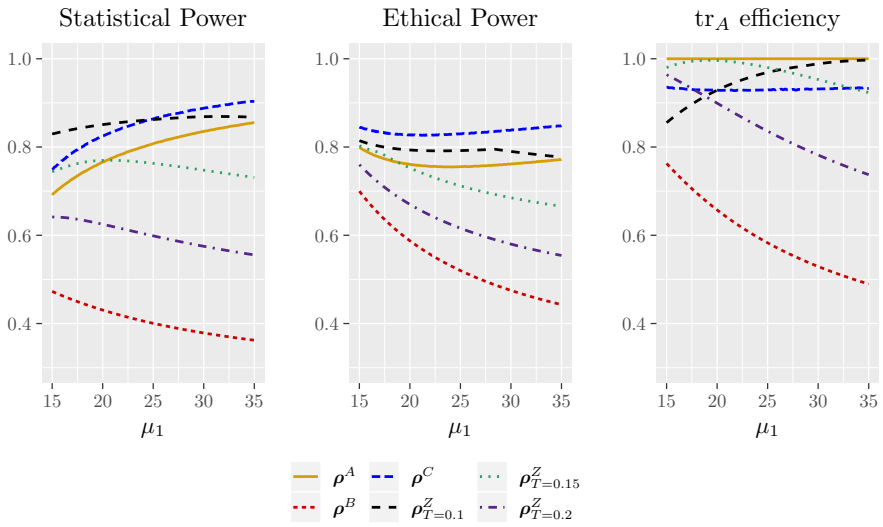

Fig. 1 Efficiency measures for $\boldsymbol{\rho}^A$, $\boldsymbol{\rho}^C$, $\boldsymbol{\rho}^B = (0.25, 0.25, 0.25, 0.25)^\top$, and $\boldsymbol{\rho}^Z$ for $T = 0.1, 0.15$ and 0.2 where $\mu_1 \in [15, 35], \mu_2 = 10, \mu_3 = 9$ and $\mu_4 = 8$.

In Table 1, we present some examples in which groups of treatments with the same efficacy are considered . We can notice that $\boldsymbol{\rho}^C$, which coincides with the balanced design in scenario *(a)*, skews the allocations to the best performing treatment as the differences between μ_1 and the other treatment effects increases. In all the configurations considered, $\boldsymbol{\rho}^C$ leads to the highest power while keeping the ethical and the estimation efficiency always greater than 92.3% and 91.5%, respectively.

Scenario	$\boldsymbol{\mu}$	$\boldsymbol{\rho}$	$E_P(\boldsymbol{\rho})$	$E_E(\boldsymbol{\rho})$	$E_{tr_A}(\boldsymbol{\rho})$
(a)	$(12, 12, 12, 10)^\top$	$\boldsymbol{\rho}^A = (0.379, 0.219, 0.219, 0.183)^\top$	0.670	0.970	1
		$\boldsymbol{\rho}^C = \boldsymbol{\rho}^B$	0.818	0.958	0.915
		$\boldsymbol{\rho}^Z$	-	-	-
(b)	$(12, 12, 10, 10)^\top$	$\boldsymbol{\rho}^A = (0.394, 0.227, 0.189, 0.189)^\top$	0.975	0.936	1
		$\boldsymbol{\rho}^C = (0.318, 0.227, 0.227, 0.227)^\top$	1	0.923	0.969
		$\boldsymbol{\rho}^B$	0.992	0.917	0.898
		$\boldsymbol{\rho}^Z$	-	-	-
(c)	$(12, 10, 10, 10)^\top$	$\boldsymbol{\rho}^A = (0.409, 0.197, 0.197, 0.197)^\top$	0.928	0.902	1
		$\boldsymbol{\rho}^C = (0.545, 0.152, 0.152, 0.152)^\top$	1	0.925	0.932
		$\boldsymbol{\rho}^B$	0.682	0.875	0.881
		$\boldsymbol{\rho}^Z$	-	-	-

Table 1 Behaviour of $\boldsymbol{\rho}^A$, $\boldsymbol{\rho}^C$ and $\boldsymbol{\rho}^B = (0.25, 0.25, 0.25, 0.25)^\top$ in presence of groups of treatments with the same efficacy.

As discussed in Section 1, it is worth noticing that the target proposed by Zhu and Hu [9] cannot be computed for some parameters configurations, e.g. scenarios *(a), (b), (c)* of Table 1. This drawback can strongly affect its applicability.

Notice also that in clinical trials comparing $K > 2$ treatments the definition of ethics is not unequivocally determined and the requirement of being ethical by skewing more patients to the superior treatment may sometimes be misleading (see [3]). The structure of the ethical constraint $\rho_i \geq \rho_{i+1}$ for $i = 1, \ldots, K-1$ in $\boldsymbol{\rho}^C$ ensures that the target has its components ordered accordingly to the magnitude of the treatment effects. In general, this property is not shared by the considered targets: for example the $\boldsymbol{\rho}^A$ allocation in (1) assigns more patients to the reference treatment, which in our set-up coincides with the best treatment. However, if for example we consider $\boldsymbol{\mu} = (10, 12, 12, 12)^\top$, then $\boldsymbol{\rho}^A = (0.325, 0.225, 0.225, 0.225)^\top$. Moreover, if we adopt $\boldsymbol{\rho}^Z$ with $T = 0.1$ in a configuration close to scenario *(a)*, e.g. $\boldsymbol{\mu} = (12.1, 12, 11.9, 10)^\top$ then the ensuing target is $\boldsymbol{\rho}^Z = (0.357, 0.1, 0.1, 0.443)^\top$. In both examples the highest proportion of patients is receiving the less effective drug showing that these targets may be inappropriate from an ethical viewpoint.

The results of Figure 1, Table 1 and the discussion above, show that the constrained optimal target $\boldsymbol{\rho}^C$ represents a valid trade-off between statistical power, inferential precision and ethical demand.

Acknowledgements Marco Novelli was 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)".

References

1. Baldi Antognini, A., Giovagnoli, A.: Adaptive designs for sequential treatment allocation. Chapman and Hall/CRC (2015)
2. Baldi Antognini, A., Novelli, M., Zagoraiou, M.: Optimal designs for testing hypothesis in multiarm clinical trials. *Stat. Meth. Med. Res.*, **28**, 3242-3259 (2019)
3. Frieri, R., Zagoraiou, M. : Optimum and ethical designs for hypothesis testing in multi-arm exponential trials. Submitted (2020)
4. Hu, F., Zhang, L. X.: Asymptotic properties of doubly adaptive biased coin designs for multitreatment clinical trials. *Ann. Stat.*, *32*, 268-301 (2004)
5. Rosenberger, W. F., Lachin, J. M.: Randomization in clinical trials: theory and practice. John Wiley and Sons (2015)
6. Silvey, S.D.: Optimal Designs. Chapman & Hall, London (1980)
7. Sverdlov, O., Rosenberger, W. F.: On recent advances in optimal allocation designs in clinical trials. *J. Stat. Theory Prac.*, **7**, 753-773 (2013)
8. Tymofyeyev, Y., Rosenberger, W. F., Hu, F.: Implementing optimal allocation in sequential binary response experiments. *J. Am. Stat. Assoc.*, **102**, 224-234 (2007)
9. Zhu, H., Hu, F.: Implementing optimal allocation for sequential continuous responses with multiple treatments. *J. Stat. Plan. Infer.*, **139**, 2420-2430 (2009)