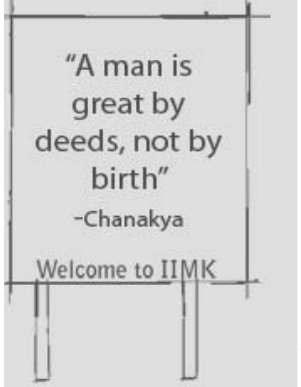

INDIAN INSTITUTE OF MANAGEMENT KOZHIKODE

Working Paper

IIMK/WPS/468/FIN/2021/08

July 2021

Patients-at-Risk (PaR): A new performance measure for response-adaptive trials

L Ramprasath[1]
Mohammed Shahid Abdulla[2]

[1]Associate Professor, Finance, Accounting and Control, Indian Institute of Management Kozhikode, IIMK Campus PO, Kunnamangalam, Kozhikode, Kerala 673 570, India; Email: lrprasath@iimk.ac.in, Phone Number (+91) 495 – 2809248

[2]Associate Professor, Information Systems, Indian Institute of Management Kozhikode, IIMK Campus PO, Kunnamangalam, Kozhikode, Kerala 673 570, India; Email: shahid@iimk.ac.in, Phone Number (+91) 495 – 2809254

Patients-at-Risk (PaR): A new performance measure for response-adaptive trials

L Ramprasath† & Mohammed Shahid Abdulla*
† *Finance, Accounting and Control Area, IIM Kozhikode*
**Information Systems Area, IIM Kozhikode*

ARTICLE HISTORY
Compiled July 7, 2021

ABSTRACT
This article aims to remove the apparent conflict between statistical power and higher allocation to the better treatment, in a particular Ethical-optimal (Et-optimal) response-adaptive design for continuous responses. An existing criterion is extended to show that the Et-optimal design could be uniformly superior over the corresponding optimal design, in finite samples. Further, one of the reasons for why experimenters prefer the standard randomized control trial over a response-adaptive trial could be the high variability of patient allocations in the latter. Though there are many response-adaptive designs in the literature which promise higher allocation to the superior treatment, this is not always assured. Here we propose a new criterion, Patients-at-Risk, for evaluating response-adaptive designs, which partly addresses this problem. Under this new criterion, an algorithm based on the explore-exploit heuristic is shown to be superior to the Et-optimal design in this particular context, thus giving a win-win solution for both ethics and statistical power.

KEYWORDS
Clinical trials; Ethics; Optimality; Response adaptive designs; Worst-case performance

1. Introduction

The main objective of a response-adaptive clinical trial design over a static randomized control trial (RCT) design is to minimize the number of patients allotted to the inferior drugs/treatments. This is easier to achieve if, based on the responses, the superior drug is identified early on during the trial. The problem of balancing both the ethical and the statistical considerations is extensively studied in the literature (see for example Rosenberger and Lachin (2015), Atkinson and Biswas (2013), Villar, Bowden, and Wason (2015) and the more recent proposal of such designs for COVID-19 related trials by Stallard et al. (2020)).

The first contribution of this article is to resolve, using a finite sample analysis, the trade-off between statistical power and ethics in two asymptotically optimal designs proposed in the literature. The first design is the one proposed in Biswas and Bhattacharya (2009) for continuous responses and the other design is its more ethical version that was proposed subsequently in Biswas and Bhattacharya (2011).

Biswas and Bhattacharya (2009) gave a simple design for sequentially assigning patients to one of two drugs, with the goal of finding the better drug. Here the re-

sponses from the two drugs correspond to two populations with unknown means μ_k and variances $\sigma_k^2, k = A, B$ and the drug with the lower μ_k is desirable. Biswas and Bhattacharya (2009) obtained their design by minimizing $N_A\psi_A + N_B\psi_B$, where N_A and N_B refer to the target sample size for the two drugs and ψ_k refers to a suitable measure of clinical optimality. This optimal design also covered the earlier designs of Biswas and Mandal (2004) and Zhang and Rosenberger (2006), as special cases. For example, in Zhang and Rosenberger (2006) the measure of clinical optimality was simply μ_k. Thus they minimized the expected trial outcome through their optimal design.

Further to this, Biswas and Bhattacharya (2011) proposed an Ethical-optimal (referred to as Et-optimal henceforth) design, where in addition to the above maximization, the design also took care of the ethical consideration that not more than a pre-fixed proportion (≤ 0.5) of patients will be allotted to the inferior drug. Because of this additional constraint, sometimes Et-optimal had lower statistical power in deciding the better drug, though it was assigning more patients to the better drug. In the articles that introduced these designs, it is also shown that both the optimal and the Et-optimal designs have allocation probabilities that converge to the optimal value, in the limit. However, the current article is concerned with their finite sample performances and a dilemma arising from such an analysis.

It is said that higher power of a design benefits future patients whereas increasing the allocation to superior treatment is useful for current patients in the study. We can also view this as another manifestation of the basic explore-exploit trade-off that is discussed in the multi-armed bandit literature (Villar, Bowden, and Wason 2015). Now, this leads to a difficult dilemma for the experimenter: Should the experimenter sacrifice the power of the clinical trial for getting a greater number of patients assigned to the better drug? The current article resolves this dilemma by extending the guideline given in Paganoni and Secchi (2007).

Though Et-optimal turns out to be superior than the optimal design according to Paganoni and Secchi (2007), it turns out that for the experimenter, the worst case performance of Et-optimal could act as a deterrent in using it in a real world setting. In order to address this, we propose a stricter criterion of Patients-at-Risk or PaR. As discussed in section 2, this could be a more useful criterion to evaluate the variability of an adaptive design than the conventional $\mathrm{Var}(N_k)$.

The article also proposes a new design, where we combine the optimal design with the explore-exploit heuristic found in bandit algorithms. This design is shown to perform better than Et-optimal design according to the more stricter PaR criterion, in finite samples. In the rest of the article, section 2 briefly reviews the optimal and Et-optimal designs and introduces the PaR criterion, while section 3 describes the new design and section 4 concludes.

2. Outperformance of Et-Optimal

We briefly state the designs of Biswas and Bhattacharya (2009) and its modification in Biswas and Bhattacharya (2011), before discussing our criterion of comparison. Consider the testing problem,

$$H_0 : \mu_B \leq \mu_A \text{ versus } H_1 : \mu_B > \mu_A,$$

where μ_A and μ_B are the means of the experimental and control drugs respectively, with the desired drug being the one with the lower mean. The optimal design in Biswas and Bhattacharya (2009) was obtained from the following optimization problem,

$$\min N_A\psi_A + N_B\psi_B$$

$$\text{subject to } \frac{\sigma_A^2}{N_A} + \frac{\sigma_B^2}{N_B} = \ell \text{ and } N_A + N_B = n.$$

Solving this yields ρ_A, the optimal targeted allocation proportion to drug A as,

$$\rho_A^{opt} = \frac{\sigma_A\sqrt{\psi_B}}{\sigma_A\sqrt{\psi_B} + \sigma_B\sqrt{\psi_A}}. \tag{1}$$

Zhang and Rosenberger (2006) for example, used a design, where for $k = A, B$, $\psi_k = \mu_k$, while Biswas and Mandal (2004) used $\psi_k = \Phi\left(\frac{\mu_k - c}{\sigma_k}\right)$, where c is a constant specified by the experimenter.

This optimal strategy was converted into an ethical-optimal strategy by Biswas and Bhattacharya (2011) by introducing the following additional constraint to the optimization problem,

$$\left(\frac{N_A}{n} - \beta\right) I(\psi_A < \psi_B) + \left(\frac{N_B}{n} - \beta\right) I(\psi_B \leq \psi_A) \geq 0$$

This constraint ensured that a minimum proportion (β) of patients is assigned to the superior drug. The new optimal solution became,

$$\rho_A = \begin{cases} \max(\rho_A^{opt}, \beta) \text{ if } \psi_A < \psi_B \\ \min(\rho_A^{opt}, 1-\beta) \text{ otherwise} \end{cases},$$

where ρ_A^{opt} is as given in (1). We refer the reader to the original papers for more details.

Throughout this article, we are assuming that the responses are instantaneous i.e., after a drug is administered, the response of a patient is available before the next patient arrives. Delayed responses can be dealt by applying all the algorithms discussed in the article only to the available responses. It is also assumed that the responses from either drug are Normal random variables, as in the original study.

We first report a comparison between the Optimal and Et-Optimal designs for a few parameter combinations in table 1 using $\psi_k = \Phi\left(\frac{\mu_k - c}{\sigma_k}\right)$. μ_A is fixed at 10, while μ_B is varied. 20,000 simulations were used for each setting, assuming a fixed trial size of $N = 150$ patients in each simulation. The first 15% of the patients in every trial were assigned randomly to either drug, to get initial estimates of the mean and variance parameters. The remaining patients were assigned as per the above allocation probabilities, with μ_k and σ_k replaced by their corresponding running estimates and $c = 10$. Throughout this article, $\beta = 0.55$ was used while implementing the Et-optimal design.

The following test statistic was used to test the null hypothesis,

$$Z = \frac{\bar{Y}_B - \bar{Y}_A}{\sqrt{\frac{\hat{\sigma}_A^2}{N_A} + \frac{\hat{\sigma}_B^2}{N_B}}},$$

where $\bar{Y}_k$ and $\hat{\sigma}_k^2$ are the end of the trial estimates of μ_k and σ_k^2, k =A, B, obtained from the responses of the patients given the corresponding drug. The test was applied at 1% level so that the difference in powers across various designs can be clearly observed.

Table 1. Power of Optimal and Et-optimal designs. Standard error for each entry is given in the bracket below the entry.

	μ_B					
	10	10.2	10.4	10.6	10.8	11
	$\sigma_A = 1,\ \sigma_B = 1$					
Optimal	0.010	0.097	0.393	0.762	0.958	0.997
	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Et-optimal	0.011	0.095	0.371	0.738	0.944	0.995
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
	$\sigma_A = 1,\ \sigma_B = 1.5$					
Optimal	0.011	0.078	0.286	0.614	0.875	0.978
	(0.001)	(0.002)	(0.003)	(0.003)	(0.002)	(0.001)
Et-optimal	0.013	0.070	0.257	0.565	0.835	0.962
	(0.001)	(0.002)	(0.003)	(0.003)	(0.003)	(0.001)
	$\sigma_A = 1.5,\ \sigma_B = 1$					
Optimal	0.011	0.060	0.213	0.464	0.738	0.914
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Et-optimal	0.012	0.060	0.199	0.456	0.724	0.909
	(0.001)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)

The critical region of the test uses the normal approximation and it is seen that the first column in table 1 justifies this approximation. Though the same values of c and the (μ, σ) combinations from the original papers are used here, it is to be noted that both the designs are scale and location invariant, when $\psi_k = \Phi\left(\frac{\mu_k - c}{\sigma_k}\right)$. For each cell in table 1, the proportion of patients assigned to drug A is given in table 2. The number of simulations for some of the cells were increased in order to bring down the standard errors and to enable the comparison in table 3, described later in this section.

Tables 1 and 2 show that the optimal design is more powerful, whereas Et-optimal design is more ethical in terms of allotting more patients to the better drug. Interestingly, this is observed across all the parameter combinations in the table. This also leads to a basic dilemma for the experimenter. Consider for example, the case of $\sigma_A = 1,\ \sigma_B = 1.5$ and $\mu_B = 10.6$. Here the optimal design has 10% higher power than Et-optimal, but Et-optimal allots 12 more patients for every 100 patients to the better drug. Now which is preferable: more powerful design or a greater number of patients to the superior treatment? Ignoring either of them can have serious consequences.

Instead of posing this as a trade-off, this dilemma is resolved using the guideline

Table 2. Allocation proportion to the better drug (A)

	μ_B				
	10.2	10.4	10.6	10.8	11
$\sigma_A = 1,\ \sigma_B = 1$					
Optimal	0.438	0.449	0.458	0.465	0.473
Et-optimal	0.506	0.527	0.537	0.541	0.542
$\sigma_A = 1,\ \sigma_B = 1.5$					
Optimal	0.386	0.394	0.402	0.408	0.414
Et-optimal	0.484	0.511	0.525	0.535	0.538
$\sigma_A = 1.5,\ \sigma_B = 1$					
Optimal	0.471	0.479	0.486	0.493	0.499
Et-optimal	0.511	0.526	0.536	0.541	0.544

given in Paganoni and Secchi (2007) with a slight modification. Let τ_1 and τ_2 be two response-adaptive designs and let $\delta = \mu_B - \mu_A$. The modified criterion is as follows. For a given δ, prefer τ_2 over τ_1 if, τ_2 is at least as powerful as τ_1 and

$$q_3(\tau_2, \delta) < q_3(\tau_1, \delta), \tag{2}$$

where $q_3(\tau, \delta)$ is the third quartile of N_B, the random variable denoting the number of patients assigned to the inferior drug when τ is used. In general, τ_2 dominates τ_1, if (2) happens for every δ, allowing the strict inequality to become an equality in some, but not all, cases.

In Paganoni and Secchi (2007), this criterion was used to compare a response-adaptive design (τ) with the static RCT design and τ would be preferred if it is at least as powerful as the RCT and $q_3(\tau, \delta) < \frac{N}{2}$. The idea of using the third quartile is to ensure that less than half the patients will be assigned to the inferior treatment with a high probability. In this paper, the modified criterion is implemented in the following manner.

Whenever Et-optimal is having lower power than optimal, the trial size for Et-optimal is gradually increased till the estimated power $\pm$ 3 std. error interval becomes equal for both the designs (up to the second decimal place). For example, for $\sigma_A = 1,\ \sigma_B = 1.5$ and $\mu_B = 11$, when the trial size was increased to 158, the 3 std. error interval for the power of the Et-optimal design became (0.97, 0.98), same as that of the optimal design. This trial size (denoted by n^*) and the corresponding third quartile of N_B is noted and compared with that of the optimal design. Table 3 gives the n^* and q_3 values. One can also use more decimal places for matching the powers, if required.

All the q_3 values reported in this and the other tables were found to be quite stable, when 20,000 simulations are done for each cell in the table. Table 3 shows that, even when a larger trial size is needed for Et-optimal to match the power of optimal, number of patients assigned to the inferior drug by the former design is possibly lower with a high probability. And this is seen to be true across all the parameter combinations considered here.

However, one cause of concern is the amount of variability in the empirical distribution of N_B. For example, in the case of $\sigma_A = 1,\ \sigma_B = 1.5$ and $\mu_B = 10.6$, even though the third quartile for Et-optimal is only 83, there was quite a bit of variability

Table 3. q_3 comparison among the two designs. Et-optimal row contains both n^* and q_3

	μ_B				
	10.2	10.4	10.6	10.8	11
	$\sigma_A = 1,\ \sigma_B = 1$				
Optimal	90	88	87	86	85
Et-optimal	150, 80	160, 80	157, 77	155, 75	150, 73
	$\sigma_A = 1,\ \sigma_B = 1.5$				
Optimal	98	96	95	94	93
Et-optimal	150, 85	163, 84	166, 83	165, 81	158, 77
	$\sigma_A = 1.5,\ \sigma_B = 1$				
Optimal	85	84	83	82	81
Et-optimal	150, 79	157, 79	152, 75	151, 74	151, 73

beyond this. In particular, the 99.5th percentile in this case was 111, which means if the clinical trial uses this design, then there is a 1-in-200 chance that more than 110 patients of the total 166 would be assigned to the inferior treatment. This defeats the main advantage of using response-adaptive designs.

Also, this variability could be one of the reasons why well-intended response-adaptive designs are not adopted in real settings. Further, the variability on the right-hand side of the distribution of N_B is the real concern and not the entire variability of N_B.

In light of the above, the following stringent modification of (2) is proposed. This is partly motivated by the idea behind the Value-at-Risk (VaR) measure that is popular in financial risk management. For a given δ, prefer τ_2 over τ_1 if,

$$\begin{aligned}&(a)\ \tau_2 \text{ is at least as powerful as } \tau_1,\\ &(b)\ q_3(\tau_2, \delta) \leq q_3(\tau_1, \delta) \text{ and}\\ &(c)\ PaR_{99.5}(\tau_2, \delta) \leq PaR_{99.5}(\tau_1, \delta),\end{aligned}$$

where PaR is the acronym for Patients-at-Risk and $PaR_{99.5}$ is the 99.5-th percentile of N_B. According to this new criterion, τ_2 dominates τ_1, if the above three conditions hold for every δ, with a strict inequality in either (b) or (c) for some δ.

The new criterion given above can also be used in the original framework of Paganoni and Secchi (2007), where a single response-adaptive design, τ, is evaluated against a static benchmark design. In such a case, τ would be preferred if in addition to $q_3(\tau, \delta) < \frac{N}{2}$, we also have, without loss of power,

$$PaR_{99.5}(\tau, \delta) < (0.5 + \epsilon)N$$

for every δ, where ϵ can be specified by the experimenter. For example, in a trial of 150 patients, keeping $\epsilon = 2\%$, would force the response-adaptive design to assign not more than 78 patients to the inferior drug with a probability of 0.995.

3. Can the PaR of Et-optimal be improved?

This section shows that it is possible to further improve the worst-case performance of Et-optimal. Based on the explore-exploit heuristic of bandit algorithms, a modification of Et-optimal is proposed and is shown to have a uniformly lower PaR, for all the parameter settings considered in section 2. We call this the Guided Play-the-Winner (GPW) design and according to this design, the trial is divided in to three stages. The allocation proportion for these stages is given in table 4.

Table 4. Stages of the GPW design

Stage	Proportion of samples	Allocation probability for drug A
1	15%	0.5
2	35%	$\hat{\rho}_A^{opt}$
3	50%	$\hat{\rho}_A^{GPW}$

The first two stages are same as the optimal design, described in section 2. In the third stage, to determine the assignment for the i-th patient, first generate $\delta_{A,i}$ from a Bernoulli($\hat{\rho}_A^{opt}$) distribution ($\delta_{A,i} = 1$ corresponds to an assignment to drug A) and the final assignment $\delta_{A,i}^{GPW}$ is determined as follows. If $\delta_{A,i} = 1$, then

$$\delta_{A,i}^{GPW} = \begin{cases} \delta_{A,i} & \text{if } (\bar{Y}_{A,i} < \bar{Y}_{B,i}) \text{ or } (N_{A,i} < N_{B,i}) \\ \text{Bernoulli}(1-\gamma) & \text{otherwise,} \end{cases}$$

where γ is the probability of playing the winner. And if $\delta_{A,i} = 0$,

$$\delta_{A,i}^{GPW} = \begin{cases} \delta_{A,i} & \text{if } (\bar{Y}_{B,i} < \bar{Y}_{A,i}) \text{ or } (N_{B,i} < N_{A,i}) \\ \text{Bernoulli}(\gamma) & \text{otherwise.} \end{cases}$$

Here $\bar{Y}_{k,i}$ and $N_{k,i}$, $k = A, B$, are the running means and the patient counts before the i-th patient arrives. Also we have implemented $\hat{\rho}_A^{opt}$ by using $c = (\bar{Y}_{A,i} + \bar{Y}_{B,i})/2$, so that the allocation proportion to drug A increases, as the gap between μ_A and μ_B widens.

The rationale for the GPW design is as follows. In the second half of the trial, the design 'exploits' (assigns the winning drug with a high probability), if Et-optimal says so. And if Et-optimal says 'explore', GPW will do it subject to the condition that the losing drug doesn't get more patients than the winner. In other words, the play-the-winner rule is "guided" by Et-optimal allocation probabilities.

We report first the power comparison between Et-optimal and GPW designs in table 5, in the same settings of our previous simulations with $\gamma = 0.95$. Table 6 gives the q_3 and $PaR_{99.5}$ values for both the designs.

Since the power of both Et-optimal and GPW designs are statistically indistinguishable, we did not increase the trial size for the design with the lower power, as we did for table 3. It is seen from tables 5 and 6 that GPW dominates Et-optimal in terms of q_3 and PaR values, across the table. To check whether this conclusion also holds for a smaller sample size, this experiment was repeated for a trial size of 80 and the results are given in tables 7 and 8.

For both $N = 150$ and $N = 80$, we did not find much difference in the results when the probability of playing the winner (γ) was varied around 0.95. From the reported

Table 5. Power of Et-optimal and GPW designs, $N = 150$

	μ_B					
	10	10.2	10.4	10.6	10.8	11
	$\sigma_A = 1,\ \sigma_B = 1$					
Et-optimal	0.011	0.095	0.371	0.738	0.944	0.995
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
GPW	0.010	0.099	0.377	0.749	0.949	0.996
	(0.001)	(0.002)	(0.003)	(0.003)	(0.002)	(0.001)
	$\sigma_A = 1,\ \sigma_B = 1.5$					
Et-optimal	0.013	0.070	0.257	0.565	0.835	0.962
	(0.001)	(0.002)	(0.003)	(0.003)	(0.003)	(0.001)
GPW	0.011	0.072	0.266	0.584	0.846	0.967
	(0.001)	(0.002)	(0.003)	(0.003)	(0.003)	(0.001)
	$\sigma_A = 1.5,\ \sigma_B = 1$					
Et-optimal	0.012	0.060	0.199	0.456	0.724	0.909
	(0.001)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)
GPW	0.012	0.060	0.206	0.462	0.735	0.912
	(0.001)	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)

Table 6. q_3 and $PaR_{99.5}$ values, $N = 150$

	μ_B				
	10.2	10.4	10.6	10.8	11
	$\sigma_A = 1,\ \sigma_B = 1$				
Et-optimal	80, 103	75, 98	74, 92	73, 87	73, 86
GPW	75, 101	74, 91	74, 75	74, 75	73, 75
	$\sigma_A = 1,\ \sigma_B = 1.5$				
Et-optimal	85, 110	78, 108	75, 103	74, 96	73, 90
GPW	75, 109	75, 101	75, 81	75, 75	74, 75
	$\sigma_A = 1.5,\ \sigma_B = 1$				
Et-optimal	79, 99	76, 96	74, 92	73, 88	73, 85
GPW	75, 98	74, 92	74, 82	73, 75	73, 75

levels in tables 5 and 7, it is observed that the normal approximation for the test statistic works better for $N = 150$, while there is a slight inflation of type-I error for $N = 80$.

It is seen from these tables that in some cases GPW is able to achieve a PaR reduction of more than 30%, without any loss of power. And this is achieved in cases which typically would have called for higher allocation to the inferior drug i.e., when the inferior drug's variability is 50% higher than the better drug. Also, the authors noted that the empirical distribution of N_B has become quite skewed under the GPW design. More particularly, it is only controlling the variability on the right-hand side of the distribution and thus addresses the typical cause of worry for the experimenter. For example, the case of $N = 80,\ \sigma_A = 1,\ \sigma_B = 1.5,\ \mu_B = 11$, with $PaR_{99.5} = 40$,

Table 7. Power of Et-optimal and GPW designs, $N = 80$

	μ_B					
	10	10.2	10.4	10.6	10.8	11
			$\sigma_A = 1,\ \sigma_B = 1$			
Et-optimal	0.014	0.063	0.199	0.446	0.710	0.901
	(0.001)	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)
GPW	0.015	0.067	0.210	0.457	0.722	0.903
	(0.001)	(0.002)	(0.003)	(0.004)	(0.003)	(0.002)
			$\sigma_A = 1,\ \sigma_B = 1.5$			
Et-optimal	0.016	0.051	0.147	0.318	0.542	0.757
	(0.001)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
GPW	0.014	0.053	0.153	0.329	0.558	0.768
	(0.001)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
			$\sigma_A = 1.5,\ \sigma_B = 1$			
Et-optimal	0.015	0.045	0.118	0.252	0.436	0.634
	(0.001)	(0.001)	(0.002)	(0.003)	(0.004)	(0.003)
GPW	0.013	0.043	0.116	0.252	0.443	0.644
	(0.001)	(0.001)	(0.002)	(0.003)	(0.004)	(0.003)

Table 8. q_3 and $PaR_{99.5}$ values, $N = 80$

	μ_B				
	10.2	10.4	10.6	10.8	11
		$\sigma_A = 1,\ \sigma_B = 1$			
Et-optimal	44, 63	42, 62	41, 60	40, 59	40, 55
GPW	40, 61	40, 58	39, 49	39, 40	39, 40
		$\sigma_A = 1,\ \sigma_B = 1.5$			
Et-optimal	47, 67	45, 66	43, 65	41, 63	41, 63
GPW	42, 65	40, 61	40, 57	40, 47	39, 40
		$\sigma_A = 1.5,\ \sigma_B = 1$			
Et-optimal	44, 62	42, 60	41, 58	40, 57	40, 56
GPW	41, 60	40, 57	39, 53	39, 48	39, 41

median ($=q_3$) = 39 and 0.5-th percentile = 21, illustrates this skew.

4. Discussion

Response-adaptive designs are quite popular among researchers working in improving the design of clinical trials. They could become more popular if their worst-case performance can also be improved along with other performance measures. This is due to the fact that the experimenter does not have the luxury of conducting multiple iterations of the same trial. The clinical trial is conducted only once. And the experimenter may not want to end up assigning, for example, 150 out of 170 patients to the

inferior treatment on a bad day. Here we have proposed the PaR criterion to partly address this issue and suggest that this should also be taken into account while comparing different adaptive designs. We also found the GPW design, which is based on the explore-exploit heuristic of bandit algorithms, to be performing uniformly better than both optimal and Et-optimal designs in terms of PaR, in all the cases considered here. In fact, the worst-case gains are as high as 30% in some cases. Besides the above contribution, the article has also removed a false dichotomy between ethics and statistical power, in the particular case of optimal and the Et-optimal designs.

References

Atkinson, A., and A. Biswas. 2013. *Randomised response-adaptive designs in clinical trials.* Boca Raton, FL: CRC Press.

Biswas, A., and R. Bhattacharya. 2009. "Optimal response-adaptive designs for normal responses." *Biometrical journal* 51:193–202.

———. 2011. "Optimal response-adaptive allocation designs in phase III clinical trials: Incorporating ethics in optimality." *Statistics and Probability Letters* 81:1155–1160.

Biswas, A., and S. Mandal. 2004. "Optimal adaptive designs in phase III clinical trials for continuous responses with covariates." In *MODa7-Advances in Model-Oriented Design and Analysis,* edited by A.Di. Bucchianico, H. Lauter, and H.P. Wynn, 51–58. Heidelberg: Physica-Verlag.

Paganoni, A.M., and P. Secchi. 2007. "A numerical study for comparing two response-adaptive designs for continuous treatment effects." *Statistical Methods and Applications* 16:321–346.

Rosenberger, W., and J. Lachin. 2015. *Randomization in clinical trials: Theory and practice.* Hoboken, NJ: John Wiley & Sons.

Stallard, N., L. Hampson, N. Benda, W. Brannath, T. Burnett, T. Friede, P.K. Kimani, et al. 2020. "Efficient Adaptive Designs for Clinical Trials of Interventions for COVID-19." *Statistics in Biopharmaceutical Research* 12(4):483–497.

Villar, S., J. Bowden, and J. Wason. 2015. "Multi-armed bandit models for the optimal design of clinical trials: Benefi ts and challenges." *Statistical Science* 30(2):199–215.

Zhang, L., and W.F. Rosenberger. 2006. "Response-adaptive randomization for clinical trials with continuous outcomes." *Biometrics* 62:562–569.

Research Office
Indian Institute of Management Kozhikode
IIMK Campus P. O.,
Kozhikode, Kerala, India,
PIN - 673 570
Phone: +91-495-2809237/ 238
Email: research@iimk.ac.in
Web: https://iimk.ac.in/faculty/publicationmenu.php